Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **521,** 3298–3322 (2023) https://doi.org/10.1093/mnras/stad587

Abell 1201: detection of an ultramassive black hole in a strong gravitational lens

J. W. Nightingale [1]★ Russell J. Smith [1] Qiuhan He [1] Conor M. O'Riordan [2] Jacob A. Kegerreis [3] Aristeidis Amvrosiadis,[1] Alastair C. Edge,[1] Amy Etherington,[1] Richard G. Hayes,[1] Ash Kelly [1] John R. Lucey [1] and Richard J. Massey [1]

[1]*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[2]*Max Planck Institute for Astrophysics, Karl-Schwarzschild-Strasse 1, D-85748 Garching bei München, Germany*
[3]*NASA Ames Research Center, Moffett Field, CA 94035, USA*

Accepted 2023 January 19. Received 2023 January 18; in original form 2022 February 28

ABSTRACT

Supermassive black holes (SMBHs) are a key catalyst of galaxy formation and evolution, leading to an observed correlation between SMBH mass $M_{\rm BH}$ and host galaxy velocity dispersion $\sigma_{\rm e}$. Outside the local Universe, measurements of $M_{\rm BH}$ are usually only possible for SMBHs in an active state: limiting sample size and introducing selection biases. Gravitational lensing makes it possible to measure the mass of non-active SMBHs. We present models of the $z = 0.169$ galaxy-scale strong lens Abell 1201. A cD galaxy in a galaxy cluster, it has sufficient 'external shear' that a magnified image of a $z = 0.451$ background galaxy is projected just ∼1 kpc from the galaxy centre. Using multiband *Hubble Space Telescope* imaging and the lens modelling software `PYAUTOLENS`, we reconstruct the distribution of mass along this line of sight. Bayesian model comparison favours a point mass with $M_{\rm BH} = 3.27 \pm 2.12 \times 10^{10}\,{\rm M}_{\odot}$ (3σ confidence limit); an ultramassive black hole. One model gives a comparable Bayesian evidence without an SMBH; however, we argue this model is nonphysical given its base assumptions. This model still provides an upper limit of $M_{\rm BH} \leq 5.3 \times 10^{10}\,{\rm M}_{\odot}$, because an SMBH above this mass deforms the lensed image ∼1 kpc from Abell 1201's centre. This builds on previous work using central images to place upper limits on $M_{\rm BH}$, but is the first to also place a lower limit and without a central image being observed. The success of this method suggests that surveys during the next decade could measure thousands more SMBH masses, and any redshift evolution of the $M_{\rm BH}$–$\sigma_{\rm e}$ relation. Results are available at https://github.com/Jammy2211/autolens_abell_1201.

Key words: gravitational lensing: strong – galaxies: evolution – galaxies: formation – quasars: supermassive black holes.

1 INTRODUCTION

Supermassive black holes (SMBHs) have emerged as an integral part of models of galaxy formation and evolution, owing to the tight correlation observed between SMBH mass, $M_{\rm BH}$, and host galaxy bulge velocity dispersion, bulge mass and other galaxy properties (Graham 2012; Kormendy & Ho 2013; van den Bosch 2016). It is posited that an SMBH resides at the centre of every galaxy and that galaxies and SMBHs coevolve with one another from their initial formation in the early Universe (Heckman & Best 2014; Smith & Bromm 2019). The mass of an individual SMBH, $M_{\rm BH}$, can be measured via spatially resolved dynamics of nearby tracers such as stars and gas (Davis et al. 2017; Thater et al. 2019). This technique has provided over 100 measurements of $M_{\rm BH}$ which show tight correlations with other galaxy properties such as bulge luminosity or velocity dispersion (Kormendy & Richstone 1995; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Graham 2001). The need for spectroscopy at high spatial resolution that resolves the SMBH's sphere of influence restricts this approach to nearby galaxies, preventing the study of how these relations evolve with redshift. Spectral fitting of active galactic nuclei (McLure & Dunlop 2004; Peterson et al. 2004; Shen 2013) and reverberation mapping techniques can provide measurements of $M_{\rm BH}$ in higher redshift galaxy populations which therefore enable evolutionary studies. However, these observations necessitate that the galaxy's SMBH is actively accreting, bringing in potential selection effects. A method that can measure $M_{\rm BH}$ for non-active galaxies outside the local Universe would be highly complementary to these existing approaches. Analysing the strong gravitational lensing of background sources, acting in some specific (and perhaps rare) circumstances and configurations, might provide such a technique.

In this paper, we present a re-examination of the strong-lensing brightest cluster galaxy (BCG) in Abell 1201. A tangential gravitational arc was first identified in shallow *Hubble Space Telescope* (*HST*) WFPC2 images of this system, by Edge et al. (2003). Compared to most cluster lenses, the arc is unusual in being formed at small projected radius from the BCG (∼2 arcsec; ∼5 kpc). Edge et al. (2003) found that a high ellipticity and/or strong external shear was necessary to match the arc shape. Integral-field spectroscopic data later revealed a faint counter-image to the main arc, projected even closer to the lens centre (∼0.3 arcsec; ∼1 kpc; Smith, Lucey &

★ E-mail: james.w.nightingale@durham.ac.uk

© 2023 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

Edge 2017a). Using a simplified position-based model of the lensing configuration, Smith et al. (2017a) argued that an additional mass of $\sim 10^{10}\,M_{\odot}$ at small radius was necessary to reproduce the counter-image as observed. The spatially resolved stellar kinematics support this conclusion (Smith, Lucey & Edge 2017b). The authors concluded that the necessary central mass could be an SMBH, but with the limited imaging data available, and the rudimentary lensing analysis employed, a degeneracy with the inner stellar mass distribution of the lens could not be excluded.

Here, we analyse new *HST* WFC3/UVIS imaging of higher spatial resolution and greater signal-to-noise ratio, using advanced lens modelling techniques, to reassess the evidence for a lensing-detected SMBH in Abell 1201. We show that the detailed structure observed in the counter image constrains the inner mass distribution of the lens, and allows us to place constraints on the central SMBH. We perform a Bayesian model comparison of a variety of lens models that include and omit a point-mass representing an SMBH. The majority of models favour the inclusion of an SMBH and produce consistent estimates of $M_{\rm BH}$, with some dependence on the form and flexibility of the assumed lens galaxy mass model. This work marks the second observation of a strong lens that provides constraints on the SMBH at the centre of its lens galaxy, following the work of Winn, Rusin & Kochanek (2004) who detected the 'central' image of a lensed source via radio observations. Our study is the first where a measurement of $M_{\rm BH}$ is inferred via strong lensing (as opposed to an upper limit) and does so without the rare observation of a central image.

Over the next decade, of order one-hundred thousand strong lenses will be discovered by cosmological surveys such as Euclid, LSST and SKA (Collett 2015), a three orders of magnitude increase over the hundreds of systems that are currently known (Bolton et al. 2008; Bolton et al. 2012; Sonnenfeld et al. 2013; Shu et al. 2016). This will naturally lead to the discovery of more exotic and peculiar strong lens systems (Orban De Xivry & Marshall 2009), whose rare lensing configurations may provide constraints on $M_{\rm BH}$. We conclude with a discussion of whether galaxy-scale strong lensing can become a viable technique to measure large samples of SMBH masses in the future[1]

This paper is structured as follows. In Section 2, we describe the *HST* imaging of Abell 1201. In Section 3, we describe the `PYAUTOLENS` method and model fits performed in this work. In Section 4, we present the results of model fits using a variety of lens models. In Section 5, we discuss the implications of our measurements, and we give a summary in Section 6. We assume a Planck 2015 cosmology throughout (Ade et al. 2016). Text files, `DYNESTY` chains and images of every model fit performed in this work are available at https://github.com/Jammy2211/autolens_abell_1201.

2 DATA

We acquired the new *HST* imaging of Abell 1201 in Programme 14886 using the Ultraviolet and VISible channel on the Wide Field Camera 3 (WFC3/UVIS). A total of five exposures with a total integration time of 7150 s were taken in the F390W bandpass, tracing the clumpy rest-frame ultraviolet emission from star-forming regions in the source galaxy. This filter probes wavelengths shorter than the 4000-Å break at the redshift of Abell 1201; hence, the foreground light of the lens is suppressed and the contrast of the source enhanced. Additionally, we acquired three exposures in F814W, totalling 1009 s, to trace the distribution of stellar mass in the BCG. The observatory-provided reduced single-exposure images were registered and combined using ASTRODRIZZLE, projecting on to an output pixel scale of 0.04 arcsec. An accurate estimate of the point-spread function (PSF) is required for the lens modelling. To this end, we employed the empirical PSF provided by STScI,[2] as appropriate to the position of the target in each individual exposure, and propagated the PSF images through the same stacking process as for the real observation. The final combined images of Abell 1201 in the two bandpasses are shown in Fig. 1.

Fig. 1 shows the F390W and F814W imaging, alongside lens-subtracted versions that highlight the lensed source galaxy. There is a giant arc 2.0–3.0 arcsec away from the lens galaxy on one side of the lens with a counter image just ∼0.3 arcsec (∼ 0.9 kpc) from the lens galaxy centre. The lens itself is a cD galaxy residing in the central regions of a galaxy cluster, in contrast to most galaxy-scale (e.g. Einstein radius <5.0 arcsec) strong lens systems that are massive elliptical field galaxies and not in a cluster environment.

The cluster Abell 1201 has also been investigated. X-ray analysis reveals an offset gas core 500 kpc northwest of the lens (Ma et al. 2012), which is interpreted as a tail of gas stripped from the offset core. The gas has a different density, entropy, and temperature than gas in the surrounding area, providing evidence indicative of a minor merger at second core passage. Alignment between the mass distribution of Abell 1201's BCG mass distribution (inferred via lens modelling performed by Edge et al. 2003) and the offset core is also noted, which could be the result of a sloshing mechanism.

3 METHOD

3.1 Overview

We use version `2022.03.30.1` of the lens modelling software `PYAUTOLENS` PYAUTOLENS (Nightingale et al. 2021a). `PYAUTOLENS` fits the lens galaxy's light and mass and the source galaxy simultaneously. The method assumes a model for the lens's foreground light (e.g. one or more Sersic profiles), which is convolved with the instrumental PSF and subtracted from the observed image. A mass model (e.g. an isothermal mass distribution) ray-traces image-pixels from the image-plane to the source-plane and a pixelized source reconstruction, using an adaptive Voronoi mesh, is performed. Fig. 2 provides an overview of a `PYAUTOLENS` lens model, where models of Abell 1201 for the image-plane lens galaxy emission and lensed source are shown alongside the source-plane source reconstruction.

By fitting the source's extended surface brightness distribution, `PYAUTOLENS` considers light rays emanating from different parts of the source; therefore, constraining different regions of the lens galaxy's potential. If a small-mass clump is near the lensed source's emission, it may cause observable distortions to one or more of its multiple images. This technique has provided detections of three non-luminous dark matter substructures (Vegetti et al. 2010, 2012; Hezaveh et al. 2016; Nightingale et al. 2022) in strong lenses, where their presence is inferred by how they perturb the appearance of the lensed images.

[1]For other lensing-related techniques, see also Event Horizon Telescope Collaboration et al. (2019) for a measurement of the SMBH at the centre of M87 by mapping the lensed shadow of surrounding gas, Banik et al. (2019) for a discussion of using strong lensing to detect intermediate-mass black holes, Chen et al. (2018) and Mahler et al. (2022) for discussions of searching for wandering SMBHs in strong lensing galaxy clusters and Hezaveh, Marshall & Blandford (2015), Tamura et al. (2015), Wong et al. (2017), Quinn et al. (2016) for studies based around strong lens central images.

[2] https://www.stsci.edu/hst/instrumentation/wfc3/data-analysis/psf

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

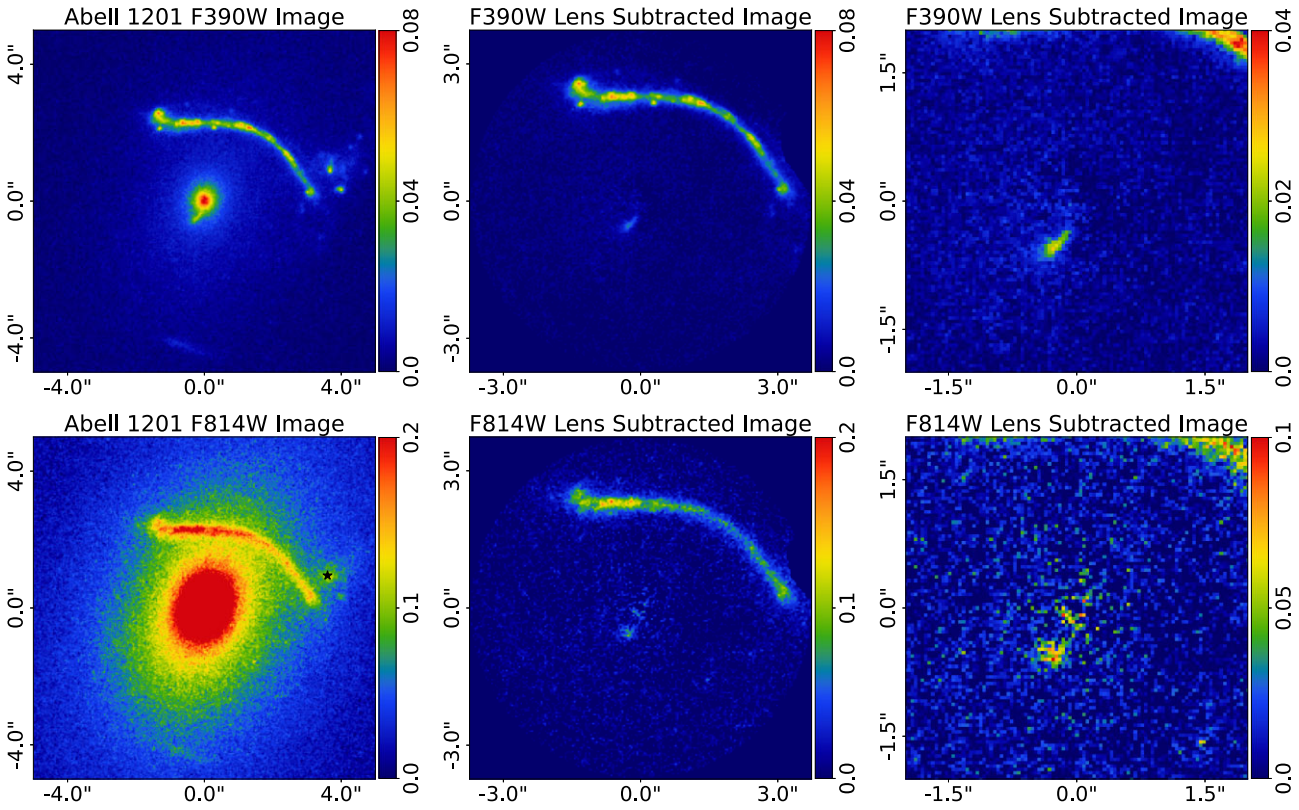

Figure 1. The observed images (left column), masked and lens subtracted images (middle column) and images zoomed in on the central regions containing the counter image (right column) of Abell 1201. The top row show the *HST* optical image taken using the F390W filter, the bottom row shows an image taken at near infrared wavelengths using the F814W filter, which are both in units of electrons per second. The lens subtractions are performed using the highest likelihood model found for each image; however, their visual appearance does not change significantly for other high likelihood models. The counter image of the giant arc can be clearly seen at both wavelengths but has much higher contrast and more clumpy structure in the bluer F390W waveband. In the F814W image, residuals from the lens light subtraction around the coordinates (0, 0) arcsec are seen; these are not a central image of the source galaxy, which would be brighter in F390W. The black star marks a line-of-sight galaxy at $z = 0.273$ which is included in certain lens models.

This work uses the same technique, albeit we are in this case investigating whether the perturbing effects of the central SMBH are detected in the lensed source emission. This is why the proximity of Abell 1201's counter-image to the lens galaxy's centre, and therefore SMBH, is so important. A high-mass SMBH will induce a local perturbation to the counter image's appearance that does not produce a significant change in the appearance of the other multiple images of source in the giant arc. This is shown in Fig. 3, where two simulated lenses based on our models of Abell 1201 are shown. In the right panel, a $M_{\rm BH} = 10^{10}\,{\rm M}_{\odot}$ SMBH is added to the lens model, which changes the location, appearance and brightness of the counter image without producing a visible change in the giant arc. Our results are therefore not based on whether the source forms a central image (Winn et al. 2004; Rusin, Keeton & Winn 2005).[3]

At the heart of the PYAUTOLENS model-fitting process is the computation of the likelihood function. We provide a brief description of this calculation in Sections 3.2–3.5. Furthermore, to assist readers less familiar with strong lens modelling, we provide Jupyter notebooks providing a visual step-by-step guide, including URL links to previous literature and explanations of technical aspects of the linear algebra and Bayesian inference. The notebooks can be found at the following link: https://github.com/Jammy2211/autolens_likelihood_function.

Recent works using PYAUTOLENS include modelling strong lenses simulated using stellar dynamics models (Cao et al. 2021) and via a cosmological simulation He et al. (2023), an automated analysis of 59 lenses (Etherington et al. 2022a, b), and studies of dark matter substructure (He et al. 2022b, a; Amorisco et al. 2022).

3.2 Coordinate system

Light and mass profile quantities are computed using elliptical coordinates $\xi = \sqrt{x^2 + y^2/q^2}$, with minor to major axis-ratio q and position angle ϕ defined counter clockwise from the positive x-axis. For model fitting, these are parametrized as two components of ellipticity

$$\epsilon_1 = \frac{1-q}{1+q}\sin 2\phi, \;\; \epsilon_2 = \frac{1-q}{1+q}\cos 2\phi. \tag{1}$$

To convert parameters from arcsecond units to stellar masses, we require the critical surface mass density

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm l}D_{\rm ls}}, \tag{2}$$

[3]By central image, we are referring to the hypothetical third or fifth image that would form directly over the centre of the lens galaxy, provided its mass distribution were sufficiently cored. We therefore do not consider the counter image located 0.3 arcsec to the southwest of the lens galaxy a central image, and will always refer to it as the counter-image.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

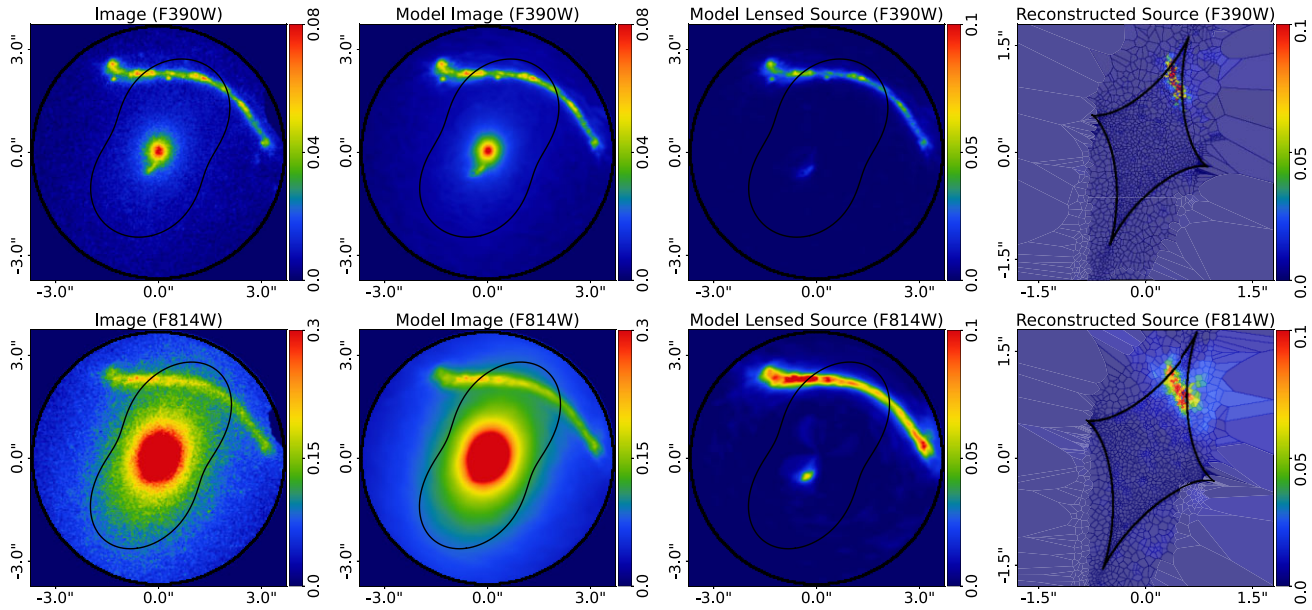

Figure 2. Fits to *HST* imaging of Abell 1201 via PYAUTOLENS. The observed data (left column), the image-plane model images of the lens and source galaxies (left-centre column), the lensed source only (right-centre column), and source-plane pixelized source reconstruction (right column) are shown. The top row shows fits to the F390W and bottom row the F814W wavebands, respectively. All images are in units of electrons per second. The lens model is the maximum likelihood model inferred at the end of the first SLaM (see Section 3.8) pipeline run, which produces a lens subtracted image. The black lines show the mass model's tangential critical curve for all panels in the image plane (central columns) and the tangential caustic for panels in the source plane (right-hand column).

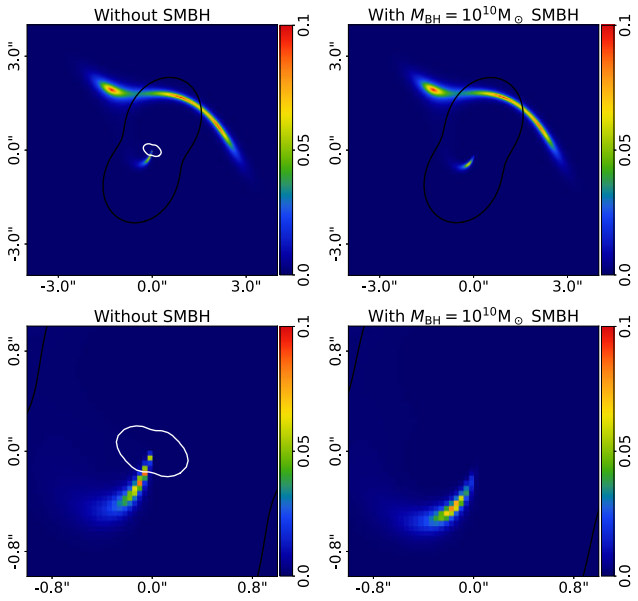

Figure 3. Illustration of how Abell 1201's lens configuration is sensitive to the lens galaxy's SMBH. Both images on the top row are simulated using the same lens mass model (a PL with external shear) and source galaxy light model (an elliptical Sersic). In the left-hand panel, an SMBH is not included, whereas in the right-hand panel an $M_{BH} = 10^{10}\,M_\odot$ SMBH is included at (0.0, 0.0), which is marked with a black cross. The bottom row shows a zoom in on the counter image. The SMBH changes the location, appearance, and brightness of the counter image but does not lead to visible changes in the giant arc. The tangential critical curve is shown by a black line and radial critical curve a white line. The latter does not form for sufficiently steep mass profiles (Kochanek 2004), including the model with an SMBH shown here.

where D_l, D_s, and D_{ls} are, respectively, the angular diameter distances to the lens, to the source, and from the lens to the source, and c is the speed of light.

3.3 Lens light models

The lens light profile intensities I are computed using one or more elliptical Sersic profiles (Sersic 1968)

$$I_{\rm Ser}(\xi_l) = I \exp\left\{-k\left[\left(\frac{\xi}{R}\right)^{\frac{1}{n}} - 1\right]\right\}, \quad (3)$$

which has seven free parameters: $(x,\ y)$, the light centre, $(\epsilon_1,\ \epsilon_2)$ the elliptical components, I, the intensity at the effective radius R and n, the Sersic index. k is a function of n (Trujillo et al. 2004). These parameters are given superscripts depending on which component of the lens galaxy they are modelling, for example the Sersic index of the bulge component is $n^{\rm bulge}$. Models with multiple light profiles are evaluated by summing each individual component's intensities. Up to three light profiles are fitted to the lens galaxy representing a bulge, bulge+disc or bulge+disc +envelope, with their superscript matching these terms. The Sersic profile intensities are computed using an adaptive oversampling routine that computes all values to a fractional accuracy of 99.99 per cent.

Bayesian model comparison is used to determine the light model complexity, from the five models listed in Table 1. These models assume one, two or three Sersic profiles and make different assumptions for how their centres and elliptical components are aligned. Model comparison is performed separately for the F390W and F814W images. Appendix A provides the priors of every Sersic profile parameter assumed in this study.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

Table 1. The five models for the lens's light that are fitted and compared in the Light pipeline. The lens light model assumes either one, two or three Sersic profiles and makes different assumptions as to whether their (x, y) centre and elliptical components (ϵ_1, ϵ_2) are aligned. A tick mark indicates that this assumption is used in the model, for example the second row is a model where both the elliptical components and centres are aligned.

Number of Sersics	Aligned elliptical components	Aligned centres
1	N/A	N/A
2	✓	✓
2	✗	✓
2	✗	✗
3	✗	✗

3.4 Lens mass models

This work fits a variety of lens galaxy mass models, which are summarized in Table 2. We fit *decomposed* mass models, where the light profile(s) that represent the lens's light are translated to stellar density profiles (via a mass-to-light profile) to perform ray-tracing (Nightingale et al. 2019).

The lens's light and stellar mass are modelled as a sum of Sersic profiles, where the Sersic profile given by equation (3) is used to give the light matter surface density profile

$$\kappa_{Ser}(\xi) = \Psi \left[\frac{q\xi}{R} \right]^{\Gamma} I_{Ser}(\xi_l) \,, \tag{4}$$

where Ψ gives the mass-to-light ratio in electrons per seconds (the units of the *HST* imaging) and Γ folds a radial dependence into the conversion of mass to light. A constant mass-to-light ratio is given for $\Gamma = 0$. If there are multiple light profile components (e.g. a bulge and disc) they assume independent values of Ψ and Γ. Deflection angles for this profile are computed via an adapted implementation of the method of Oguri (2021), which decomposes the convergence profile into multiple cored steep elliptical profiles and efficiently computes the deflection angles from each.

Observationally, early-type galaxies are observed to exhibit steep internal gradients in some spectral features associated with dwarf stars. If these features are truly driven by variations in the initial mass function, as advocated by van Dokkum et al. (2017), La Barbera et al. (2019), then substantial stellar mass-to-light ratio gradients are expected (Ferreras et al. 2019). Some evidence for such trends have indeed been reported by previous lensing studies using decomposed mass models (e.g. Oldham & Auger 2018). We therefore fit a stellar mass model which allows for different mass-to-light ratios and radial gradients in each stellar component (bulge, disc, and envelope). This ensures that we do not incorrectly favour the inclusion of an SMBH, as could otherwise occur if there is no other way for the lens model to increase the amount of mass centrally.

The dark matter component is given by an elliptical Navarro–Frenk–White (NFW) profile. Parameters associated with the lens's dark matter have superscript 'dark'. The NFW represents the universal density profile predicted for dark matter halos by cosmological *N*-body simulations (Zhao 1996; Navarro, Frenk & White 1997) and with a volume mass density given by

$$\rho = \frac{\rho_{\rm s}^{\rm dark}}{(r/r_{\rm s}^{\rm dark})(1 + r/r_{\rm s}^{\rm dark})^2} . \tag{5}$$

The halo normalization is given by $\rho_{\rm s}^{\rm dark}$ and the scale radius by $r_{\rm s}^{\rm dark}$. The dark matter normalization is parametrized using the mass at 200 times the critical density of the Universe, $M_{200}^{\rm dark}$, as a free parameter. The scale radius is set via $M_{200}^{\rm dark}$ using the mean of the mass–concentration relation of Ludlow et al. (2016), which uses the lens and source redshifts to convert this to units of solar masses.

The dark matter model has five free parameters: $(x^{\rm dark}, y^{\rm dark})$, the centre, $(\epsilon_1^{\rm dark}, \epsilon_2^{\rm dark})$, the elliptical components and; the mass, $M_{200}^{\rm dark}$. In Appendix D, we fit an elliptical NFW using a parametrization which also varies the concentration of the NFW, to test models which can increase the dark matter central density. The deflection angles of the elliptical NFW are computed via the same method used for the Sersic profile (Oguri 2021).

An external shear field is included and parametrized as two elliptical components $(\epsilon_1^{\rm ext}, \epsilon_2^{\rm ext})$, where parameters associated with the lens's external shear have superscript 'ext'. The shear magnitude $\gamma^{\rm ext}$ and the orientation of the semimajor axis $\theta^{\rm ext}$, measured counter-clockwise from north, are given by

$$\gamma^{\rm ext} = \sqrt{\epsilon_1^{\rm ext^2} + \epsilon_2^{\rm ext^2}}, \ \tan 2\phi^{\rm ext} = \frac{\epsilon_2^{\rm ext}}{\epsilon_1^{\rm ext}} . \tag{6}$$

Deflection angles are computed analytically.

To test for the presence of an SMBH via Bayesian model comparison, every model is fitted with and without a point-mass, whose parameters have superscript 'smbh'. This model includes a

Table 2. The light and mass profiles used in this work. Column 1 gives the model name. Column 2 whether it models the lens's light, mass, or both. Column 3 states what component of mass it represents. Column 4 gives its associated parameters and units.

Model	Component	Represents	Parameters	
Point mass	Mass	Black hole	$\theta_{\rm E}^{\rm smbh}$: Einstein radius (arcsec)	
Sersic	Light+ Mass	Stellar matter (Bulge, disc, envelope)	(x, y): centre (arcsec) I: intensity n: Sersic index Ψ: mass-to-light ratio ($e^- s^{-1}$)	(ϵ_1, ϵ_2): elliptical components R: effective radius (arcsec) Γ: radial gradient
Elliptical NFW	Mass	Dark matter	$(x^{\rm dark}, y^{\rm dark})$: centre (arcsec) $M_{200}^{\rm dark}$: mass at 200 ($M_\odot$)	$(\epsilon_1^{\rm dark}, \epsilon_2^{\rm dark})$: elliptical components
Shear	Mass	Line of sight	$(\epsilon_1^{\rm ext}, \epsilon_2^{\rm ext})$: elliptical components	
Elliptical Power law (PL)	Mass	Total (stellar + dark matter)	$(x^{\rm mass}, y^{\rm mass})$: centre (arcsec) $\theta_{\rm E}^{\rm mass}$: Einstein radius (arcsec)	$(\epsilon_1^{\rm mass}, \epsilon_2^{\rm mass})$: elliptical components $\gamma^{\rm mass}$: density slope
Broken Power law (BPL)	Mass	Total (stellar + dark matter)	$(x^{\rm mass}, y^{\rm mass})$: centre (arcsec) $\theta_{\rm E}^{\rm mass}$: Einstein radius (arcsec) $r_{\rm B}^{\rm mass}$: break radius (arcsec)	$(\epsilon_1^{\rm mass}, \epsilon_2^{\rm mass})$: elliptical components $t_1^{\rm mass}$: inner density slope $t_2^{\rm mass}$: outer density slope

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

single free parameter, the Einstein radius $\theta_{\rm Ein}^{\rm smbh}$, which is related to mass as

$$M_{\rm BH} = \Sigma_{\rm crit}\, \pi\, (\theta_{\rm Ein}^{\rm smbh})^2. \quad (7)$$

$\theta_{\rm Ein}^{\rm smbh}$ is in units of arcseconds and $M_{\rm BH}$ in stellar masses. Point mass deflection angles are computed analytically. The SMBH ($x^{\rm smbh}$, $y^{\rm smbh}$) centre is aligned with the highest Sersic index light profile (e.g. the bulge) for decomposed mass models.

In Appendix D, we fit *total* mass models that represent all the mass (e.g. stellar plus dark) in a single profile, either the elliptical power law (PL; Tessore, Bellagamba & Metcalf 2016) or the elliptical broken power law (BPL) introduced by O'Riordan, Warren & Mortlock (2019), O'Riordan, Warren & Mortlock (2020), and O'Riordan, Warren & Mortlock (2021). Parameters associated with the total mass model have superscript 'mass'. For these models the SMBH ($x^{\rm smbh}$, $y^{\rm smbh}$) centre is aligned with the centre of the PL or BPL mass profile. The results of fitting this model are summarized in the main paper.

We fit a number of additional lens mass models that make different assumptions, in order to verify that none change any of this paper's main results. An additional galaxy is present towards the right of the giant arc, as shown in the first panel of Fig. 1. In Appendix E, we include this galaxy in the lens mass model. In Appendix F, we fit models where the SMBH position is free to vary. In Appendix G, we fit mass models with a shallower inner density profile, which form a large radial critical curve.

3.5 Source reconstruction

After subtracting the foreground lens emission and ray-tracing the coordinates to the source-plane via the mass model, the source is reconstructed in the source-plane using an adaptive Voronoi mesh which accounts for irregular or asymmetric source morphologies (see Fig. 2). Our results use the `PYAUTOLENS` pixelization `VoronoiBrightnessImage`, which adapts the centres of the Voronoi pixels to the reconstructed source morphology, such that more resolution is dedicated to its brighter central regions (see Nightingale, Dye & Massey 2018).

The reconstruction computes the linear superposition of PSF-smeared source pixel images that best fits the observed image. This uses the matrix $f_{\rm ij}$, which maps the jth pixel of each lensed image to each source pixel i. When constructing $f_{\rm ij}$ we apply image-plane subgridding of degree 4×4, meaning that $16 \times j$ subpixels are fractionally mapped to source pixels with a weighting of $\frac{1}{16}$, removing aliasing effects (Nightingale & Dye 2015).

Following the formalism of (Warren & Dye 2003, WD03 hereafter), we define the data vector $\vec{D}_i = \sum_{\rm j=1}^{J} f_{ij}(d_j - b_j)/\sigma_j^2$ and curvature matrix $\mathsf{F}_{ik} = \sum_{\rm j=1}^{J} f_{ij} f_{kj}/\sigma_j^2$, where d_j are the observed image flux values with statistical uncertainties $\sigma_{\rm j}$, and b_j are the model lens light values. The source pixel surface brightnesses values are given by $s = \mathsf{F}^{-1}\vec{D}$, which are solved via a linear inversion that minimizes

$$\chi^2 = \sum_{\rm j=1}^{J} \left[\frac{(\sum_{\rm i=1}^{I} s_i f_{ij}) + b_j - d_j}{\sigma_j} \right]. \quad (8)$$

The term $\sum_{\rm i=1}^{I} s_i f_{ij}$ maps the reconstructed source back to the image plane for comparison with the observed data.

This matrix inversion is ill-posed; therefore, to avoid overfitting noise, the solution is regularized using a linear regularization matrix H (see WD03). The matrix H applies a prior on the source reconstruction, penalizing solutions where the difference in reconstructed flux of neighbouring Voronoi source pixels is large. We use the `PYAUTOLENS` regularization scheme `AdaptiveBrightness`, which adapts the degree of smoothing to the reconstructed source's luminous emission (see Nightingale et al. 2018). The degree of smoothing is chosen objectively using the Bayesian formalism introduced by Suyu et al. (2006). The likelihood function used in this work is taken from Dye et al. (2008) and is given by

$$-2\ln\mathcal{L} = \chi^2 + s^T\mathsf{H}s + \ln\left[\det(\mathsf{F}+\mathsf{H})\right] - \ln\left[\det(\mathsf{H})\right] + \sum_{\rm j=1}^{J} \ln\left[2\pi(\sigma_j)^2\right]. \quad (9)$$

The step-by-step Jupyter notebooks linked to above describes how the different terms in this likelihood function compare and ranks different source reconstructions, allowing one to objectively determine the lens model that provides the best fit to the data in a Bayesian context.

3.6 Data preparation

In both, the F390W and F814W wavebands there is emission from nearby interloper galaxies towards the right of the giant arc, which can be most clearly seen in the upper left-hand panel of Fig. 1. Including this emission would negatively impact our analysis; therefore, we remove it beforehand. Our lens analysis assumes a circular mask of radius 3.7 arcsec, whereby all image-pixels outside this circular region are not included in the fitting procedure. The central panels of Fig. 1 show that this mask removes the majority of foreground emission; however, a small fraction is still within this circle. We therefore subtract it using a graphical user interface, replacing it with background noise in the image and increasing the RMS noise-map values of these pixels to ensure they do not contribute to the likelihood function. We also consider lens models that include this galaxy in the ray-tracing (see Appendix E).

3.7 Light model waveband

The wavelength at which the lens galaxy's emission is observed is important for tracing its stellar mass distribution. The F390W image of Abell 1201 observes the lens galaxy at rest-frame ultra-violet wavelengths, possibly probing younger stellar populations with lower mass-to-light ratios. The F814W image observes rest-frame near infrared (NIR) emission and probes more aged and reddened stellar populations that make up a greater fraction of the stellar mass. This can be seen in Fig. 1, where only the central regions of the lens are visible in the F390W image compared to the F814W image. The F390W image is therefore less appropriate for constraining the stellar mass component of the lens model.

Therefore, to fit the decomposed mass model to the F390W image we use the maximum likelihood Sersic light model parameters of the F814W fits that are chosen after the lens light Bayesian model comparison (see Section 3.3). The mass-to-light ratio and gradient parameters of each Sersic remain free to vary, ensuring high flexibility in the model's stellar mass distribution. Fits are performed using a lens light subtracted image for the F390W image which is output midway through the analysis. To ease the comparison between fits to the F390W and F814W images, we follow the same approach with the F814W image, using the same fixed maximum likelihood Sersic parameters and fitting a lens subtracted image output midway through the analysis.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

3.8 SLaM pipelines

The models of lens mass, lens light and source light are complex and their parameter spaces highly dimensional. Without human intervention or careful set up, a model-fitting algorithm (e.g. a Markov chain Monte Carlo search) may converge very slowly to the global maximum likelihood solution. PYAUTOLENS therefore applies 'non-linear search chaining' to break the search into a sequence of tractable operations. Using the probabilistic programming language PyAutoFit PyAutoFit, we fit a series of parametric lens models that approximate the form of the desired model, with growing complexity. A fit to the simplest model provides information to initialize a fit to the next model. The final search is started around the global maximum likelihood and with priors reflecting the likelihood surface. Each fit in this chain uses the nested sampler (DYNESTY Speagle 2020) joshspeagle/DYNESTY. The models used to perform this analysis extend the Source, Light and Mass (SLaM) pipelines described by Etherington et al. (2022a, hereafter E22), Cao et al. (2021), and He et al. (2023). They are available at https://github.com/Jammy2211/autolens_workspace.

The first pipeline, called the Source pipeline, initializes the pixelized source model by inferring a robust lens light subtraction (using a double Sersic model) and total mass model (using a PL with $\gamma = 2$ plus shear). The highest likelihood lens model and source reconstruction at this stage of the pipeline are shown in Fig. 2. They give an accurate foreground lens subtraction and reconstruction of the lensed source's light.

The Light pipeline follows, which uses fixed values of the mass and source parameters corresponding to the maximum likelihood model of the Source pipeline. The lens's mass is therefore again fitted using a total mass model such that the lens light model does not yet contribute to the ray-tracing. The only free parameters in this pipeline are those of the lens light and all five of the models listed in Table 1 are fitted independently, enabling Bayesian model comparison. The results of the Light pipeline, including the models chosen for all subsequent model fits, are presented in Appendix B.

The final pipeline is the Mass pipeline, which in E22 directly follows the Light Pipeline, fitting PL mass profiles representing the total mass distribution. In this work, we do not immediately start the Mass pipeline after the Light pipeline, due to the complications of fitting the stellar component of the decomposed models to the F390W imaging data discussed previously (see Section 3.7). Instead, the lens light models favoured by model comparison are used to output lens-subtracted F390W and F814W images. An analysis of these images is then performed from scratch, starting a new SLaM pipeline fit that uses a scaled down Source pipeline, which removes models that fit the lens light, and which omits the Light pipeline completely (see E22).

When this analysis reaches the Mass pipeline, it fits the decomposed models (models assuming two or three Sersic profiles for the lens light and stellar mass) and the total mass models (the PL and BPL), whose results are described in Appendix D. Every mass model is fitted twice, with and without a point mass representing an SMBH. The Bayesian model comparison of these mass models is the main component of this work's results.

As described in E22, the SLaM pipelines use prior passing to initialize the regions of parameter space that DYNESTY will search in later pipelines, based on the results of earlier pipelines. Appendix A gives a description of the priors used in this work. We also use the likelihood cap analysis described in E22 to estimate errors on lens model parameters.

3.9 Bayesian evidences

The Bayesian evidence, $\mathcal{Z}$, of every lens model we fit is estimated by DYNESTY and is given by equation (2) of Speagle (2020). The Bayesian evidence is the integral over all parameters in the model and therefore naturally includes a penalty term for including too much complexity in a model – if a model has more free parameters it is penalized for this complexity. The evidence is computed via sampling of equation (9). Our analysis therefore incorporates the principle of Occam's razor, whereby more complex models are only favoured if they improve the fit enough to justify their additional complexity compared a simpler model. To compare models, we use the difference in log evidence, $\Delta \ln \mathcal{Z}$. An increase of $\Delta \ln \mathcal{Z} = 4.5$ for one model over another corresponds to odds of 90:1 in favour of that model. For comparisons of lens models with and without an SMBH this corresponds to a 3σ detection of the SMBH. An increase of $\Delta \ln \mathcal{Z} = 11$ corresponds to a 5σ detection.

However, there are sources of uncertainty in the evidence estimate that means taking these numbers at face value is problematic. For example, there is an error on the evidence estimated by DYNESTY, with identical runs of a lens model showing variations of $\ln \mathcal{Z} \sim 5$ (due to stochasticity in the DYNESTY sampling process). Adjusting the priors on the lens model parameters or reparametrizing the model also change its value, with tests showing variations up to $\ln \mathcal{Z} \sim 5$. Accordingly, we consider values of $\Delta \ln \mathcal{Z} > 10$ sufficient to favour more complex models over simpler ones, including the detection of an SMBH.

4 RESULTS

We now present the results of lens modelling of Abell 1201. We first examine the preferred choice of lens light models, inferred using an isothermal mass model that omits an SMBH. Then we present results using the more complex stellar plus dark matter decomposed mass models, which may also include an SMBH. We discuss additional mass models which assume a total mass profile. In each case, we examine the reconstruction of the near-centre counter image that is highly sensitive to the central mass distribution and SMBH, as well as the quantitative Bayesian evidence, $\mathcal{Z}$.

4.1 Lens Light model

The choice of lens light model via Bayesian model comparison is described in Appendix B and summarized as follows:

(i) All light models with two or three Sersic profiles are favoured over models with one Sersic, producing Bayesian evidence increases of $\Delta \ln \mathcal{Z} > 300$.

(ii) The two Sersic models whose centres, position angles, and axis ratios are unaligned produce values of $\Delta \ln \mathcal{Z} > 100$ compared to two Sersic models that assume alignment.

(iii) For the F814W image, the three Sersic models marginally give the highest evidence overall, where $\Delta \ln \mathcal{Z} = 12$ compared with the double Sersic with unaligned geometric parameters. We use this image to create the lens light subtracted image that mass models are fitted to.

(iv) The triple Sersic could not be constrained in the F390W band, owing to its observed lower rest-frame wavelength. We therefore use the double Sersic with aligned parameters to create the F390W lens light subtracted image.

Fig. 4 shows the highest evidence lens light model fits to the F390W and F814W images. A good fit to the lens galaxy's emission

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

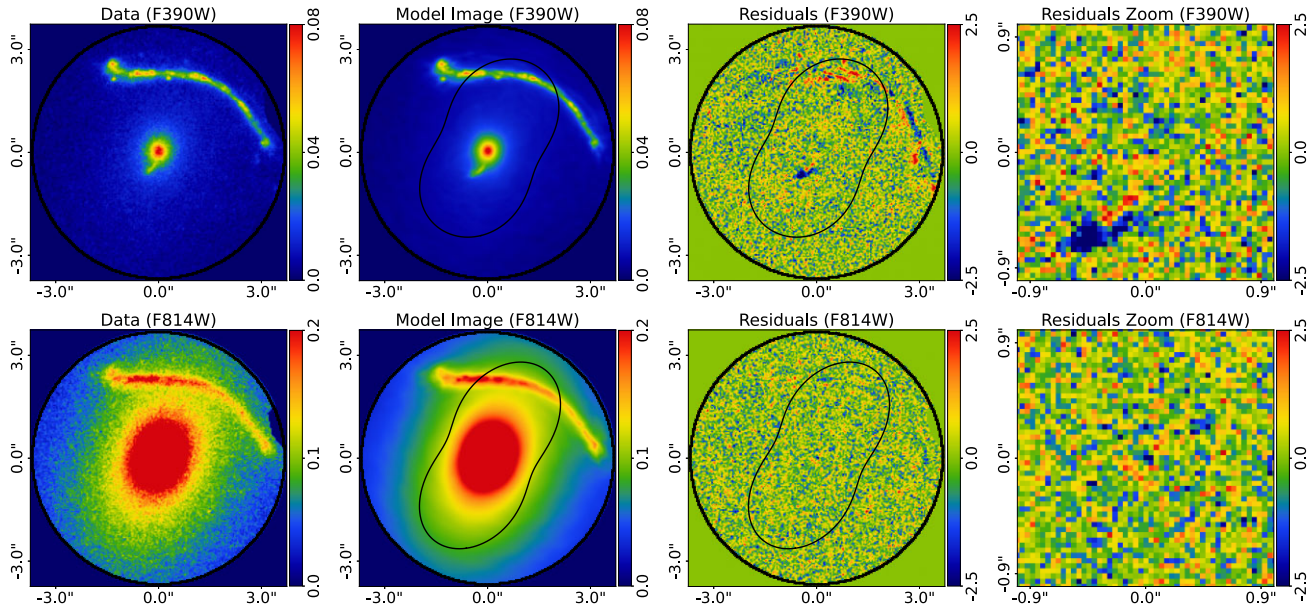

Figure 4. Lens light subtractions of *HST* imaging of Abell 1201. The observed images (left column), image-plane model images of the lens and source galaxies (left-centre column), normalized residuals (right-centre column) and a zoom-in of these residuals near the counter image (right column) are shown. The top row shows the F390W and bottom row the F814W wavebands respectively. The lens model is the maximum likelihood model inferred at the end of the first SLaM pipeline run, which produces a lens subtracted image. For the F390W data, a double Sersic lens light model with offset centres and elliptical components is shown, whereas for the F814W data a triple Sersic model is used. The magenta circle indicates residuals that are due to the lens light subtraction. The counter image is fitted poorly in the F390W image, because the mass model (which is an isothermal mass model with shear) does not enable an accurate reconstruction of the lensed source's structure.

Table 3. The inferred model parameters of the lens light models with one, two and three Sersic profiles fitted to the F814W image in the Light pipeline. The two Sersic model does not assume alignment in its geometric parameters. Errors are given at 3σ confidence intervals.

Model	Component	x (arcsec)	y (arcsec)	ϵ_1	ϵ_2	I ($e^- s^{-1}$)	$R_{\rm eff}$ (arcsec)	n
Sersic x1	Bulge	$0.001^{+0.003}_{-0.003}$	$0.008^{+0.003}_{-0.003}$	$0.087^{+0.005}_{-0.005}$	$-0.086^{+0.005}_{-0.005}$	$0.010^{+0.002}_{-0.002}$	$7.18^{+1.28}_{-1.08}$	$3.78^{+0.18}_{-0.18}$
Sersic x2	Bulge	$-0.009^{+0.003}_{-0.004}$	$0.003^{+0.003}_{-0.003}$	$0.030^{+0.011}_{-0.012}$	$-0.062^{+0.010}_{-0.001}$	$0.26^{+0.01}_{-0.02}$	$0.46^{+0.04}_{-0.04}$	$1.25^{+0.08}_{-0.07}$
	Disc	$0.069^{+0.015}_{-0.012}$	$0.032^{+0.018}_{-0.016}$	$0.16^{+0.015}_{-0.012}$	$-0.14^{+0.011}_{-0.013}$	$0.030^{+0.005}_{-0.006}$	$5.14^{+0.94}_{-0.63}$	$1.31^{+0.26}_{-0.12}$
Sersic x3	Bulge	$-0.005^{+0.004}_{-0.005}$	$-0.002^{+0.004}_{-0.004}$	$0.047^{+0.010}_{-0.013}$	$-0.046^{+0.015}_{-0.013}$	$0.22^{+0.012}_{-0.014}$	$0.46^{+0.03}_{-0.03}$	$1.28^{+0.06}_{-0.06}$
	Disc	$0.12^{+0.06}_{-0.04}$	$-0.048^{+0.045}_{-0.0657}$	$0.22^{+0.042}_{-0.030}$	$-0.11^{+0.03}_{-0.02}$	$0.025^{+0.004}_{-0.006}$	$4.63^{+1.42}_{-1.30}$	$1.16^{+0.25}_{-0.28}$
	Envelope	$-0.071^{+0.079}_{-0.06}$	$0.033^{+0.08}_{-0.10}$	$0.032^{+0.079}_{-0.083}$	$-0.27^{+0.061}_{-0.068}$	$0.0024^{+0.0042}_{-0.0018}$	$12.15^{+17.13}_{-7.71}$	$2.37^{+0.90}_{-0.64}$

and a clean subtraction is seen. Table 3 gives a subset of inferred parameters for fits to the F814W data and the full results of lens light model comparison are presented in Appendix B.

A small magenta circle is plotted on this figure and subsequent figures to indicate where the centre of the lens galaxy is. Within this magenta circle faint correlated residuals due to a slightly imperfect lens light subtraction can be seen. These are more visible in the F814W image, which is expected given the lens stellar emission is much brighter. The residuals appear as a dipole-like feature, which is commonly seen for lens light subtractions of *HST* imaging of strong lenses (e.g. Etherington et al. 2022a).

We considered whether these residuals might be a central image of the lensed source galaxy, but in this case the feature would be much brighter in the F390W image. Dust absorption could lower the F390W emission, however *HST* F606W observations of Abell 1201 also show no central emission Smith et al. (2017a), making dust absorption unlikely. In Appendix G we fit mass models with priors manually tuned to include a centrally cored mass profile, which for the F814W (or the F390W) data do not reconstruct this central emission. Lens modelling therefore confirms it is not a central image.

4.2 Decomposed mass models

We now present results using decomposed mass models that separately model Abell 1201's stellar and dark matter. Based on the lens light model comparison, we fit models assuming both two and three Sersic profiles (where parameters for the F390W fits use those inferred from fits to the F814W, see Section 3.7). We fit both models independently to both the F390W and F814W images. Visualization in this section shows the triple Sersic fits, Appendix C shows figures for the double Sersic fits.

Table 4 shows the values of $\ln \mathcal{Z}$ for decomposed models with and without an SMBH. Values of $\Delta \ln \mathcal{Z} > 30$ are seen for all model fits to the F390W image with an SMBH compared to those without. The highest overall value of $\ln \mathcal{Z}$ is the triple Sersic decomposed mass model with an SMBH, which is a $\ln \mathcal{Z}$ value more than 60

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

Table 4. The Bayesian Evidence, $\ln \mathcal{Z}$, of each model fit performed by the Mass pipelines using decomposed mass models that assume two and three Sersic profiles, an elliptical NFW and external shear. Fits to both the F390W and F814W images are shown, where the F390W fits assume the Sersic parameters of the F814W image for the stellar mass. The favoured model given our criteria of $\Delta \ln \mathcal{Z} > 10$ is shown in bold. For the F390W image, all models with an SMBH produce $\Delta \ln \mathcal{Z}$ values of at least 30 above models without an SMBH.

Filter	Number of Sersics	Includes SMBH?	$\ln \mathcal{Z}$
F390W	2	✗	125 637.18
F390W	2	✓	125 669.13
F390W	3	✗	125 598.48
F390W	3	✓	**125 699.06**
F814W	2	✗	78 330.51
F814W	2	✓	78 328.12
F814W	3	✗	78 329.19
F814W	3	✓	78 332.19

greater than that for any decomposed model without an SMBH. For the F814W images all models produce nearly consistent values of $\ln \mathcal{Z}$ with or without an SMBH, indicating that the higher S/N of the F390W data or the source's different structure is enabling the SMBH detection.

Fig. 5 shows the reconstructed lensed sources and normalized residuals for fits to the F390W and F814W images with and without an SMBH. All models reproduce the giant arc and counter image. Residuals are seen around the giant arc in the F390W image indicating missing complexity in the mass model. These residuals are seen across all mass models compared in this work (including fits which include the mass of the line-of-sight galaxy to the right of the giant arc, see Appendix E). We therefore do not anticipate they impact our inference on the SMBH. The reconstructed counter images for the models with and without an SMBH are visibly distinct and they produce different residuals, albeit this is difficult to discern from Fig. 5 due to the large arc-second scales over which the image is plotted.

Fig. 6 therefore shows zoom-ins around the counter image for the F390W image, where models without and with an SMBH are shown on the top and bottom rows respectively. The model without an SMBH places extraneous flux in the reconstructed counter image, which is not present when the SMBH is included (this flux can be seen within the radial critical curve shown by a white line and the magenta circle, and is not related to the lens light residuals). Fig. 7 shows zoom-ins for the F814W image where the same extraneous flux is seen for the model not including an SMBH. The same behaviour is seen in Appendix C for the decomposed models which fits two Sersics instead of three. The inclusion of the SMBH therefore allows the counter image to be reconstructed more accurately in both wavebands, removing central luminous emission that is not observed in the data. Removing this extraneous flux increases $\ln \mathcal{Z}$ for the F390W image data only, implying that the F814W data are too low S/N for changes in the counter image reconstruction to improve the fit in a Bayesian sense.

The inferred one-dimensional (1D) convergence profiles for the decomposed models with three Sersic profiles are shown in Fig. 8. When an SMBH is included the inferred mass model convergence is shallower. Increasing the central density of the lens galaxy's mass model therefore produces a similar lensing effect to including an SMBH and is an alternative way to improve the counter image fit. We will expand on this further when we discuss the total mass model fits.

Fig. 9 shows the two-dimensional (2D) probability density functions (PDFs) of the mass model parameters and the SMBH normalization $\theta_{\rm Ein}^{\rm smbh}$ for fits to the F390W image using the triple Sersic decomposed models. $\theta_{\rm Ein}^{\rm smbh}$ depends on the parameters controlling the mass distribution (e.g. $\Gamma^{\rm bulge}$, $\Gamma^{\rm disc}$, $\Gamma^{\rm env}$).

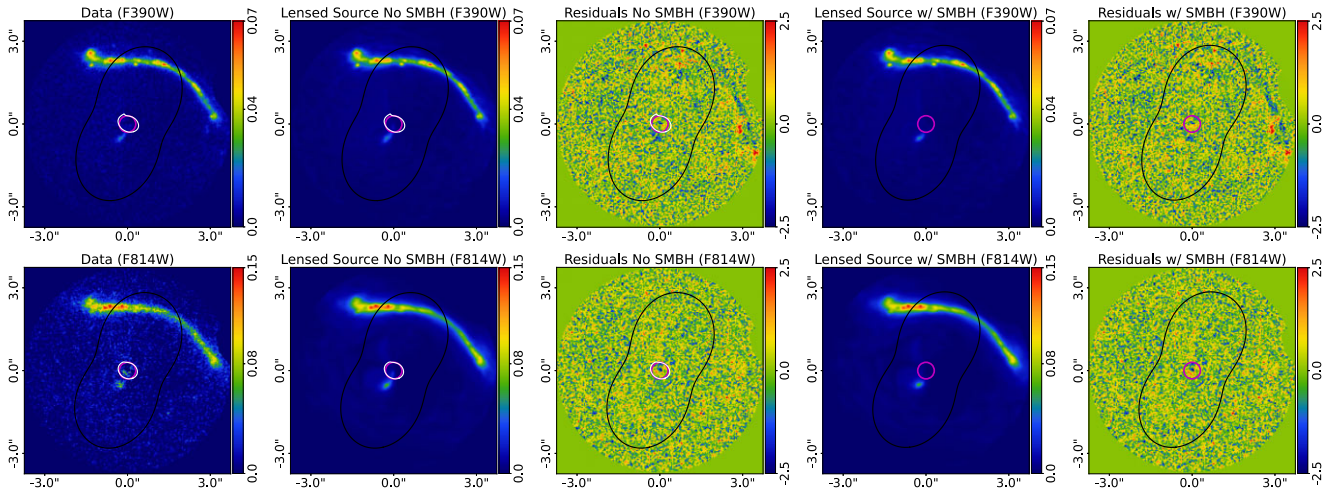

Figure 5. The observed image (left panel), model lensed source and normalized residuals for decomposed model fits without an SMBH (left-centre and centre panels) and with an SMBH (right-centre and right panel). The top row shows fits to the F390W image and bottom row the F814W image. Minimal residuals are seen in the central regions, indicating the lens light model and subtraction are accurate (the lens light models are visualized in Fig. B1). The magenta circle marks regions of the data where the brightest regions of the lens light were observed and subtracted. The tangential caustic is shown by a black line and radial critical curve and caustic a white line; the latter does not form for models including an SMBH.
Source reconstructions with and without an SMBH successfully reproduce the giant arc and counter image, although residuals are present in both indicating their detailed structure is not fitted accurately. Figs 6 and 7 show zoom-ins of the counter image to better illustrate how these different models reconstruct the data.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

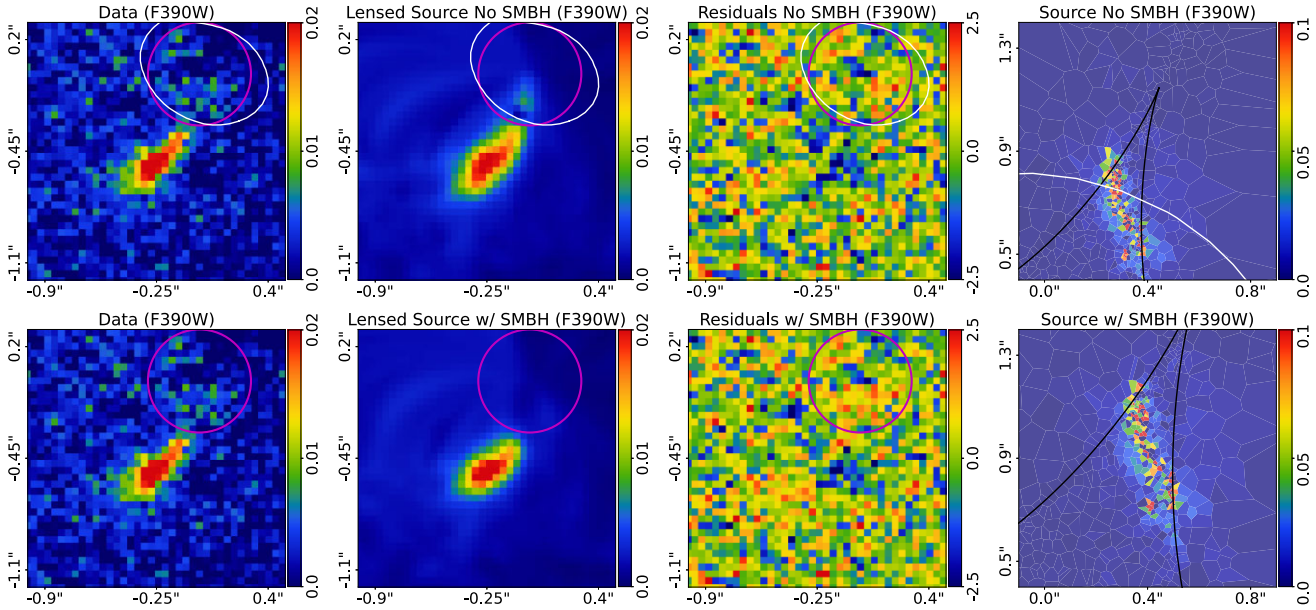

Figure 6. Zoom-ins of the observed counter image in the F390W data (left panel), the model lensed source (left-centre panel), the normalized residuals (right-centre panel) and the source reconstruction (right panel). The top and bottom rows show triple Sersic plus NFW decomposed model fits without and with an SMBH respectively. All models include an external shear. The magenta circle marks regions of the data where the brightest regions of the lens light were observed and subtracted. Models which omit an SMBH form extraneous light in the reconstructed counter image (which is seen just inside the magenta circle), which is not present in the data. The tangential caustic is shown by a black line and the radial critical curve and caustic are shown with a white line; the latter does not form for models including an SMBH.

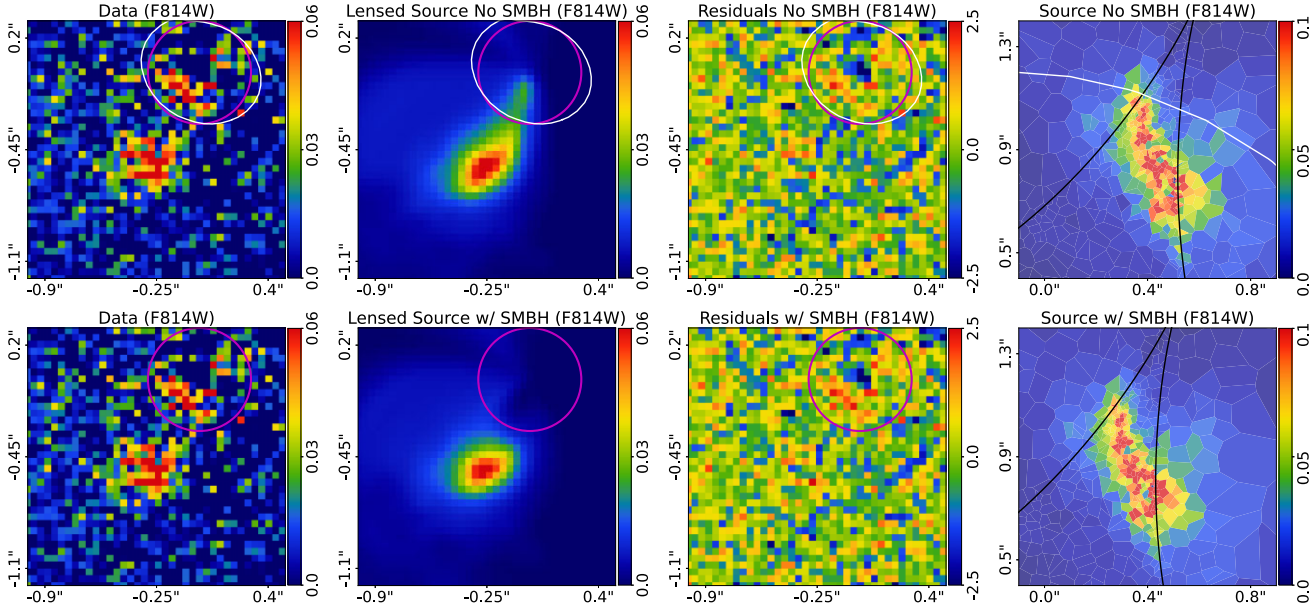

Figure 7. The same as Fig. 6 but for the F814W data.

Tables 5 and 6 give the inferred parameter estimates of the decomposed models. We can compare our inferred dark matter halo mass to the virial mass estimate of Rines et al. (2013) (from the infall caustic method), which for an NFW dark matter halo gives $M_{200} = 3.9 \pm 0.1 \times 10^{14}\mathrm{M}_{\odot}$. Table 6 shows our estimates of M_{200} range between $M_{200} = 0 - 5 \times 10^{14} M_{\odot}$. Both models with an SMBH are consistent with Rines et al. (2013). Our lens model is therefore inferring a realistic dark matter host halo.

4.3 Total mass models

The results of fitting total mass models that represent the stars and dark matter with a single projected mass distribution are given in Appendix D. For the PL mass model, which has reduced flexibility in adjusting its central density, the inferred $\ln \mathcal{Z}$ values without an SMBH are over 100 below models including an SMBH (PL or decomposed). When the PL includes an SMBH, $\ln \mathcal{Z}$ increases

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

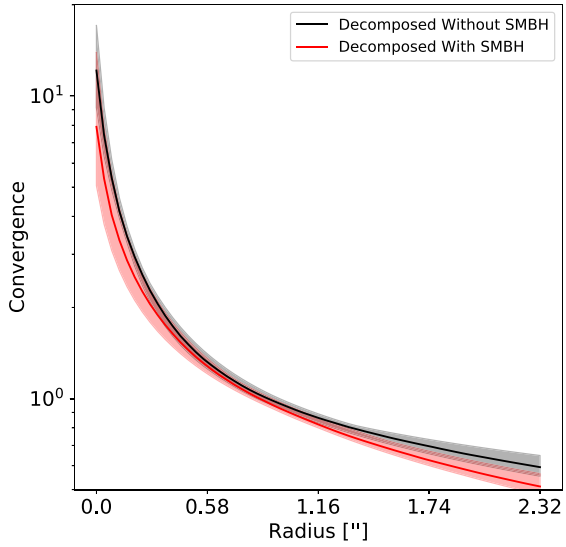

Figure 8. The convergence as a function of radius inferred using the F390W image for the decomposed mass models which assume three Sersic profiles, without an SMBH (black) and with an SMBH (red). All models include an external shear. Each line is computed using coordinates that extend radially outwards from the centre of the mass profile and are aligned with its major axis. Shaded regions for each mass model's convergence are shown, corresponding to the inferred 3σ confidence intervals. The 1D convergence of the SMBH is not included, to make comparison of each mass model's convergence straightforward.

to within ∼10 of the decomposed models with an SMBH. The PL fits therefore strongly favour an SMBH. The counter image reconstructions also reflect those seen above, whereby PL models without an SMBH show extraneous flux which is removed when an SMBH is included.

For the BPL, which has more flexibility in adjusting its central density, the model without an SMBH infers $\ln \mathcal{Z} = 125699.90$ for the F390W data. This is within 10 of the highest evidence decomposed and PL models with an SMBH. This model also reconstructs the counter image without extraneous flux. In a Bayesian sense, the BPL model without an SMBH is therefore as likely as any model fitted in this work with an SMBH, calling into question whether an SMBH is necessary in the lens model.

The inferred BPL model increases its central density above any decomposed model inferred in Section 4.2. In Appendix D, we therefore verify that the decomposed model parametrization can attain the same central density as the BPL. We show that it does when the bulge's radial gradient parameter is increased to $\Gamma^{\rm bulge} = 0.9$. The reason we do not infer this model is because this model is lower likelihood than models inferred above, where $\Gamma^{\rm bulge} = \sim 0.5$, indicating that increasing the stellar mass density produces a different lensing effect to including an SMBH. In Appendix D, we also fit models where the dark matter concentration is free to vary, such that it can reach the same central density as the BPL. These models again do not produce solutions with as high an evidence as those presented above.

For the high-density BPL model to fit the data as well as the mass models with an SMBH, its $(x^{\rm mass}, y^{\rm mass})$ centre assumes values that are ≥100 pc offset from the centre of the bulge's luminous emission. In Appendix D, we show that if the BPL model centre is aligned with the luminous bulge it produces a much lower $\ln \mathcal{Z}$. The BPL model is built-on the assumption that it can simultaneously represent both the stellar and dark matter mass distributions (O'Riordan et al. 2019). Therefore, on the grounds that a 100 pc offset between the light and total mass distribution is non-physical and breaks the underlying assumption on which the BPL is built, we favour models including an SMBH which do not require this offset.

4.4 Alternative models

We verify that the inclusion of an SMBH is still favoured for a number of alternative lens galaxy mass models. In Appendix E, we include the ray-tracing effects of the line-of-sight galaxy to the north-east of the giant arc, by modelling it as a singular isothermal sphere. In Appendix F, we fit lens models which allow the centre of the SMBH to vary as a free parameter. In Appendix G, we explore a family of solutions where the lens mass model has a shallow (or cored) inner density, therefore forming a larger radial critical curve than those inferred in the main paper. For all alternative models, an SMBH is favoured with the same or greater significance than shown for the models above.

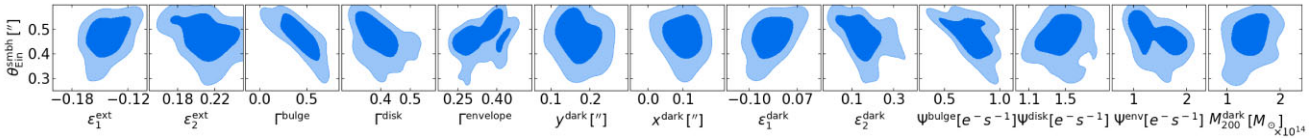

Figure 9. The 2D PDF of fits to the F390W image of Abell 1201 using the triple Sersic decomposed mass model. Marginalized 2D contours are shown for every lens mass model parameter paired with the SMBH normalization $\theta_{\rm Ein}^{\rm smbh}$ which is related to $M_{\rm BH}$, see equation (7). The inner and outer contours cover the 1 and 2σ confidence intervals respectively.

Table 5. The inferred stellar mass, shear, and SMBH model parameters of the decomposed mass models with two and three Sersic components fitted to the F390W image in the mass pipeline. Errors are given at 3σ confidence intervals.

Model	$\Psi^{\rm bulge}$ ($e^- s^{-1}$)	$\Psi^{\rm disc}$ ($e^- s^{-1}$)	$\Psi^{\rm envelope}$ ($e^- s^{-1}$)	$\Gamma^{\rm bulge}$	$\Gamma^{\rm disc}$	$\Gamma^{\rm envelope}$	$\epsilon_1^{\rm ext}$	$\epsilon_2^{\rm ext}$	$\theta_{\rm Ein}^{\rm smbh}(")$
x3 Sersic	$1.47^{+0.69}_{-0.55}$	$1.14^{+0.44}_{-0.28}$	$2.32^{+1.05}_{-0.95}$	$0.57^{+0.21}_{-0.32}$	$0.34^{+0.19}_{-0.18}$	$0.12^{+0.08}_{-0.12}$	$-0.097^{+0.042}_{-0.044}$	$0.19^{+0.05}_{-0.07}$	
x2 Sersic	$1.65^{+0.69}_{-0.61}$	$0.95^{+0.74}_{-0.30}$		$0.46^{+0.35}_{-0.34}$	$0.29^{+0.16}_{-0.14}$		$-0.084^{+0.061}_{-0.057}$	$0.13^{+0.06}_{-0.06}$	
x3 Sersic + SMBH	$0.54^{+0.26}_{-0.38}$	$1.45^{+0.55}_{-0.33}$	$1.69^{+0.63}_{-1.05}$	$0.50^{+0.37}_{-0.25}$	$0.41^{+0.14}_{-0.11}$	$0.33^{+0.20}_{-0.12}$	$-0.14^{+0.02}_{-0.02}$	$0.22^{+0.02}_{-0.04}$	$0.48^{+0.07}_{-0.11}$
x2 Sersic + SMBH	$0.87^{+0.61}_{-0.80}$	$1.25^{+0.57}_{-0.46}$		$0.15^{+0.65}_{-1.12}$	$0.48^{+0.12}_{-0.18}$		$-0.14^{+0.04}_{-0.03}$	$0.20^{+0.03}_{-0.04}$	$0.42^{+0.09}_{-0.17}$

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

Table 6. The inferred dark matter model parameters of the decomposed mass models with two and three Sersic components fitted to the F390W image in the Mass pipeline. Errors are given at 3σ confidence intervals.

Model	$x^{\rm dark}$ (arcsec)	$y^{\rm dark}$ (arcsec)	$\epsilon_1^{\rm dark}$	$\epsilon_2^{\rm dark}$	$M_{200}^{\rm dark}$ ($M_\odot \times 10^{14}$)
x3 Sersic	$0.048^{+0.108}_{-0.126}$	$0.050^{+0.100}_{-0.121}$	$0.022^{+0.110}_{-0.090}$	$0.13^{+0.12}_{-0.15}$	$3.43^{+1.24}_{-7.76}$
x2 Sersic	$0.031^{+0.059}_{-0.058}$	$0.11^{+0.17}_{-0.12}$	$0.032^{+0.084}_{-0.106}$	$0.017^{+0.119}_{-0.097}$	$5.55^{+3.29}_{-2.19}$
x3 Sersic + SMBH	$0.13^{+0.08}_{-0.17}$	$0.17^{+0.03}_{-0.10}$	$-0.036^{+0.066}_{-0.115}$	$0.19^{+0.12}_{-0.16}$	$1.43^{+6.48}_{-6.28}$
x2 Sersic + SMBH	$-0.22^{+0.12}_{-0.12}$	$0.008^{+0.084}_{-0.088}$	$-0.015^{+0.086}_{-0.159}$	$0.095^{+0.094}_{-0.067}$	$1.69^{+8.31}_{-4.37}$

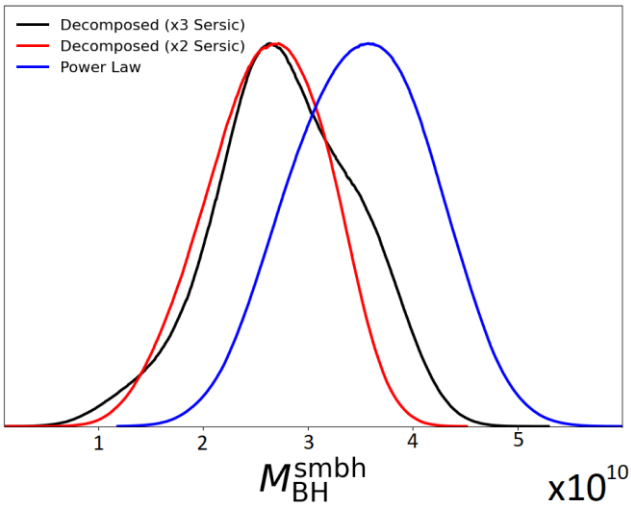

Figure 10. The 1D PDFs of the SMBH mass $M_{\rm BH}$ for fits to the F390W image of Abell 1201. Inferred values of $M_{\rm BH}$ are shown for the decomposed mass model with three Sersics, two Sersics, and the PL total mass model. The BPL fitted in Appendix D and discarded due to its non-physical 100 pc offset.

4.5 SMBH mass

The 1D PDFs for $M_{\rm BH}$ for the decomposed two and three Sersic models and PL total mass model are shown in Fig. 10. At 3σ confidence, the SMBH mass inferred for fits to the F390W image for different mass models (excluding the BPL due to its non-physical offset centre) are as follows :

(i) $M_{\rm BH} = 2.22^{+1.41}_{-1.06} \times 10^{10}$ M$_\odot$ for the triple Sersic decomposed model.
(ii) $M_{\rm BH} = 2.91^{+1.17}_{-0.91} \times 10^{10}$ M$_\odot$ for double Sersic fits.
(iii) $M_{\rm BH} = 3.95^{+1.44}_{-1.47} \times 10^{10}$ M$_\odot$ for the PL model.

To estimate a final value of $M_{\rm BH}$, we simply estimate the value which spans the full range of measurements, producing $M_{\rm BH} = 3.27 \pm 2.12 \times 10^{10}$ M$_\odot$ at 3σ confidence.

4.6 Upper limit analysis

Although we have discarded the BPL model on the grounds of physical plausibility, it can still be used to place an upper limit on $M_{\rm BH}$, even with the offset centre. Once an SMBH of sufficiently high mass is included in the mass model, it deforms the counter image reconstruction in a way which cannot be compensated for by reducing the inner density of the mass model. To demonstrate this, Fig. 11 shows the reconstructed counter images of a BPL model fit without an SMBH and with an SMBH whose mass is fixed to $M_{\rm BH} = 10^{11}$ M$_\odot$. The SMBH displaces the counter image, producing a reconstruction that is not consistent with the observed data.

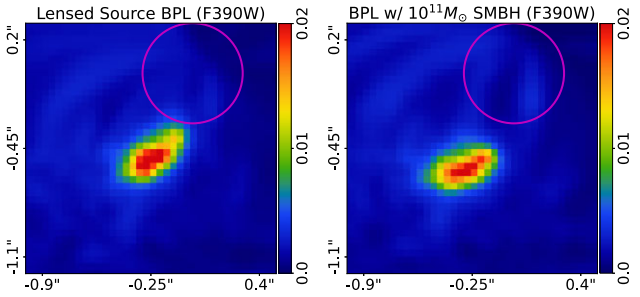

Figure 11. The reconstructed counter image for a BPL models without an SMBH (left-hand panel) and including a $M_{\rm BH} = 1.0 \times 10^{11}$ M$_\odot$ SMBH for fits to the F390W image. The fit including a $M_{\rm BH} = 10^{11}$ M$_\odot$ SMBH displaces the reconstructed counter image such that it does not fit the data accurately.

Table 7. The Bayesian evidences, ln $\mathcal{Z}$, of BPL mass model fits that include an SMBH with a fixed mass. A 1D grid of fits are shown, which iteratively increase the SMBH mass $M_{\rm BH}$. For SMBHs above masses of $M_{\rm BH} = 5.349 \times 10^{10}$ M$_\odot$, all ln $\mathcal{Z}$ values are at least 20 below the BPL model without an SMBH where ln $\mathcal{Z} = 125699.90$. Therefore, SMBHs above this mass are ruled out by the data, because they deform the reconstruction of the counter image (see Fig. 11).

$\theta_{\rm Ein}^{\rm smbh}$ ($''$)	$M_{\rm BH}$ (M$_\odot \times 10^{10}$)	ln $\mathcal{Z}$
None	0.0	125699.90
0.2	0.513	125706.18
0.3	1.145	125693.19
0.4	2.030	125686.63
0.5	3.168	125657.25
0.6	4.557	125676.59
0.625	4.945	125699.68
0.65	5.349	**125655.86**
0.675	5.765	125676.04
0.7	6.202	125636.15
0.725	6.651	125624.91
0.75	7.118	125648.55
0.775	7.580	125656.45
0.8	8.099	125617.61
0.9	10.248	125464.21

The value $M_{\rm BH} = 10^{11}$ M$_\odot$ was chosen to visually emphasize how a high-mass SMBH disfigures the counter image. We can fit a grid of BPL plus SMBH models where $M_{\rm BH}$ is fixed to incrementally higher values between 1 and 10 $\times$ 10^{10} M$_\odot$ to the F390W data. Table 7 shows the ln $\mathcal{Z}$ values for each fit, where a drop of ln $\mathcal{Z} = 20$ is seen above masses of $M_{\rm BH} = 5.3 \times 10^{10}$ M$_\odot$. The BPL model with a non-physical offset centre therefore still places an upper limit of $M_{\rm BH} \leq 5.3 \times 10^{10}$ M$_\odot$.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

While in this study Abell 1201's counter image contains sufficient information to provide a measurement of $M_{\rm BH}$, in less fortuitous circumstances upper limits on $M_{\rm BH}$ will still be possible in many strong lenses.

5 DISCUSSION

5.1 Supermassive black holes

5.1.1 $M_{BH}-\sigma_e$ *relation*

Fig. 12 shows the inferred value of $M_{\rm BH} = 3.27 \pm 2.12 \times 10^{10}\,{\rm M}_{\odot}$ on the black-hole mass versus velocity dispersion relation. This figure shows that Abell 1201 has one of the largest reported black hole masses measured so far, making it an ultramassive black hole (Hlavacek-Larrondo et al. 2012). Its mass is comparable to the SMBH of the brightest cluster galaxies NGC 3842 and NGC 4889 (McConnell & Ma 2013) and the field elliptical NGC 1600 (Thomas et al. 2016), all of which are measured via stellar orbit analysis. All three objects have similar values of $\sigma_{\rm e}$ to Abell 1201.

The SMBH of Abell 1201 is a $\sim 2\sigma$ outlier above the scatter of the $M_{\rm BH}-\sigma_{\rm e}$ relation. Two other objects with similar $\sigma_{\rm e}$ values to Abell 1201, NGC3842, and NGC1600 are ~ 1.5–2σ outliers above the mean relation. There are no corresponding outliers at $\sim 1.5\sigma$ below the mean relation, indicating that for $\sigma_{\rm e} > 250\,{\rm km\,s^{-1}}$ SMBH masses tend to be above the mean $M_{\rm BH}-\sigma_{\rm e}$ relation. Although there are too few objects to draw definitive conclusions, such an upwards kink at high $\sigma_{\rm e}$ is a prediction of different physical processes. For example, binary SMBH scouring, which saturates $\sigma_{\rm e}$ whilst increasing $M_{\rm BH}$ (Kormendy & Bender 2013; Thomas et al. 2014), as well as AGN feedback processes (Hlavacek-Larrondo et al. 2012).

5.1.2 Stellar core

Massive ellipticals are often observed with a stellar core, quantified via the Nuker or cored Sersic models (Hernquist 1990; Trujillo et al. 2004; Dullo & Graham 2013, 2014). BCGs like Abell 1201 may have extremely large and flat cores (Postman et al. 2012). It is posited that these cores form via SMBH scouring, whereby the dissipationless merging of two SMBHs in the centre of a galaxy preferentially ejects high-mass stars via three-body interactions (Faber et al. 1996; Merritt 2006; Kormendy et al. 2009; Kormendy & Bender 2013; Thomas et al. 2014). We fitted the core-Sersic model to Abell 1201's lens galaxy light during our initial analysis; however, the model did not produce an improved fit to the data. Typical core sizes are 0.02−0.5 kpc (Dullo 2019); therefore, if Abell 1201 has a stellar core it may be we simply cannot resolve it, due to the data's resolution of $\sim 120\,{\rm pc\,pixel^{-1}}$.

Aspects of the lens models that include an SMBH point towards a cored (or at least shallow) inner density. For example, the PL mass model with an SMBH infers a slope $\gamma^{\rm mass} = 1.65^{+0.12}_{-0.12}$, which is much shallower than many massive elliptical strong lenses with near isothermal slopes of $\gamma^{\rm mass} \sim 2$ (Koopmans et al. 2009). Decomposed models including an SMBH give comparable inner densities. When fitting the core-Sersic model we only included it in the model for the lens galaxy's light. We did not fit it as part of a decomposed mass model and therefore did not try to constrain the stellar core via the ray-tracing and lensing analysis. Future studies hunting for SMBHs in strong lenses may benefit from doing this, because an improved model of the lens's central mass density could help break the degeneracy seen in this work with $M_{\rm BH}$.

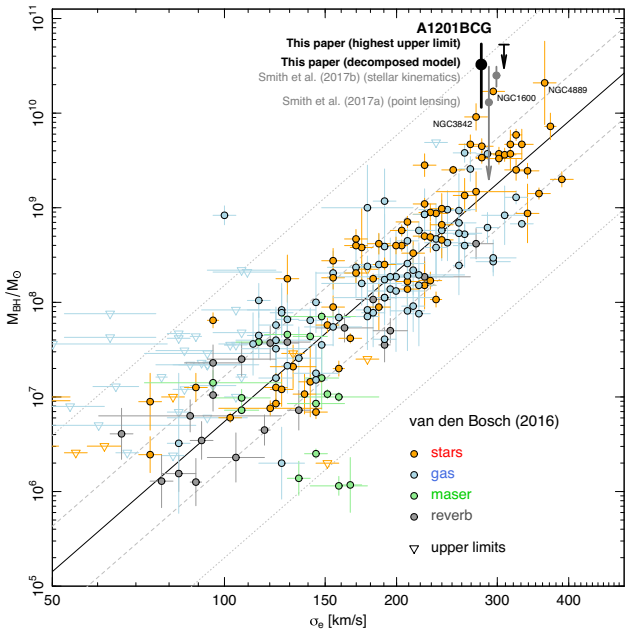

Figure 12. This work's measurements of Abell 1201's SMBH's mass in comparison to the black-hole mass versus velocity dispersion relation, from the compilation of van den Bosch (2016). Abell 1201's $\sigma_{\rm e}$ value is taken from Smith et al. (2017b). This work's measurement of $M_{\rm BH} = 3.27 \pm 2.12 \times 10^{10}\,{\rm M}_{\odot}$ is shown in black, which comes from averaging over all mass models. The upper limit of $M_{\rm BH} \leq 5.3 \times 10^{10}\,{\rm M}_{\odot}$ inferred for the BPL mass model (without an SMBH) is shown for completeness, although we have argued this model is less trustworthy due to being non-physical (see Appendix D). This figure is adapted from Smith et al. (2017a) and shows their inferred SMBH masses in grey, which come from independent analyses using either point-source based strong lens modelling (Smith et al. 2017a) or stellar kinematics (Smith et al. 2017b). Both works report that an SMBH with $M_{\rm BH} \geq 10^{10}\,{\rm M}_{\odot}$ fits the data, but neither work could break a degeneracy with models that assumed a radial gradient in the conversion of mass to light. Our inferred value of $M_{\rm BH}$ in Abell 1201 makes it one of the highest mass SMBH's measured. The grey dashed and dotted diagonal lines show 1σ and 2σ scatter of the mean $M_{\rm BH}-\sigma_{\rm e}$ relation, with Abell 1201's SMBH approximately a 2σ positive outlier.

5.1.3 Outlook for strong lensing

Abell 1201 is the second strong lens in which the central SMBH mass has been constrained. It is the first to do so without a central image, as well as the first to provide a measurement of $M_{\rm BH}$ as opposed to an upper limit. This raises a number of questions: what is so special about Abell 1201 that makes it sensitive to its SMBH? Can $M_{\rm BH}$ be measured in other known strong lenses? How common an occurrence will this be amongst the incoming samples of 100 000 strong lenses?

Abell 1201 is a unique strong lens in that its counter image is close to the lens centre and it is a cD galaxy in a galaxy cluster. The cluster potential exerts a large external shear (which is seen in our lens models) that brings the counter-image even closer to the lens centre (Smith et al. 2017a), an effect that is not present in most known galaxy-scale strong lenses, which are typically field galaxies. Thus, a very specific set of circumstances may make Abell 1201 sensitive to its SMBH, and a strategy to finding more systems is to target cD / BCG galaxies with instruments like Multi Unit Spectroscopic Explorer (MUSE).

On the other hand, some known strong lenses in surveys like the Sloan Lens ACS Survey (Bolton et al. 2008) and Strong Lensing in the Legacy Survey (Sonnenfeld et al. 2013) may be sensitive to their

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

central SMBH and appropriate lens modelling has simply not been performed. Certainly, every strong lens will provide an upper limit on $M_{\rm BH}$, the question is whether any are low enough to be informative for models of galaxy evolution. Whilst the multiple images of strong lenses are predominantly observed at radii well beyond Abell 1201's 1 kpc counter image, there are examples of strong lenses where the extended emission of the lensed source goes this close. For example, SLACS1250 + 0523, which was modelled by Nightingale et al. (2019). In many surveys, for a candidate strong lens to be worthy of following up with higher resolution imaging, a visible counter image clearly distinct from the lens's emission is typically required. Systems like Abell 1201 may therefore be common in nature but rarely selected for follow-up. We leave it to future work to investigate what constraints known strong lenses can place on $M_{\rm BH}$.

It has long been expected that strong lensing can constrain SMBH masses when a central third or fifth image is observed (Mao, Witt & Koopmans 2001; Rusin & Ma 2001; Keeton 2003; Hezaveh et al. 2015). Such a system was presented by Winn et al. (2004), who placed an upper limit of $M_{\rm BH} < 2 \times 10^8\,{\rm M}_\odot$. These systems require the inner density profile of the lens galaxy to be sufficiently cored that the central image is not demagnifed below the observing instrument's detection limit. Given that no other such observation has been made despite numerous attempts (Jackson et al. 2015; Wong et al. 2017), this appears to be a rare occurrence. Lower limits on $M_{\rm BH}$ have been placed in systems, where a central image is not detected (Quinn et al. 2016).

Abell 1201 demonstrates that an SMBH mass measurement is possible even when the lens's inner density is not cored. This offers hope that large samples of strong lenses can one day constrain the $M_{\rm BH}-\sigma_{\rm e}$ relation. This would enable the masses of non-active black holes to be measured at high redshifts, and would provide measurements on the high $\sigma_{\rm e}$ end of the relation where few ETGs are observed in the local Universe. With over 100 000 strong lenses set to be observed in the next decade (Collett 2015), it is inevitable that more SMBH measurements via strong lensing will be made; however, more work is necessary to determine how common an occurrence this will be, and in what types of strong lenses and at how high of a redshift such constraints are feasible. If the detectability of a strong lens's SMBH depends on a specific set of circumstances like Abell 1201, there will also be unavoidable selection effects that must be accounted for.

6 SUMMARY

We present an analysis of the galaxy-scale strong gravitational lens Abell 1201 using multiwaveband *HST* imaging. Abell 1201 is a particularly unique system for two reasons: (i) its counter image is just 1 kpc away from the lens galaxy centre; (ii) it is a cD galaxy located within a galaxy cluster. After extensive strong lens modelling, we show that reconstructions of Abell 1201's counter image provide constraints for mass models that include a central SMBH. After performing Bayesian model comparison, we find that all but one lens model of Abell 1201 prefer the inclusion of an SMBH. By averaging over these models, we infer a value of $M_{\rm BH} = 3.27 \pm 2.12 \times 10^{10}\,{\rm M}_\odot$, in agreement with previous lensing and stellar dynamics models of the system Smith et al. (2017a, b). This makes it one of the largest black hole masses measured to date and qualifies it as an ultramassive black hole. Its mass is comparable to other high-velocity dispersion $\sigma_{\rm e}$ systems on the $M_{\rm BH}-\sigma_{\rm e}$ relation whose masses were measured via stellar orbit analysis.

There is one mass model without an SMBH which, in a Bayesian sense, is as plausible as any model including an SMBH. This model has a lot of flexibility in adjusting its central density and mimics the lensing effect of the SMBH by increasing its density to be extremely peaked; far more so than any other mass model. However, the model simultaneously requires that its mass centre is offset from the luminous centre of the bulge by $\geq$100 pc. This offset is not necessary when an SMBH is included in the lens model and we therefore rule-out this model as being non-physical. Even allowing for this non-physical offset centre, the model still provides an upper limit of $M_{\rm BH} \leq 5.3 \times 10^{10}\,{\rm M}_\odot$, as including an SMBH above this mass completely deforms the counter image reconstruction. Therefore, even strong lens systems which are not as fortuitous as Abell 1201 in their configuration could provide meaningful constraints on SMBHs as upper limits.

Abell 1201 is the second strong lens to provide constraints on its central SMBH mass, following the upper limit of $M_{\rm BH} \leq 2 \times 10^8\,{\rm M}_\odot$ placed by Winn et al. (2004) in a strong lens whose central image was observed. Our work is therefore the first to not only place an upper limit but measure $M_{\rm BH}$ and it does so without the rare observation of a central image. This offers hope that many more strong lens systems can potentially constrain the mass of their central SMBH, although the unique properties of Abell 1201 may mean this remains a somewhat rare occurrence. Further investigation is necessary to draw firm conclusions, but with over one hundred thousand strong lens systems set to be discovered in the next decade there is hope that strong lensing can one day constrain the redshift evolution of the $M_{\rm BH}-\sigma_{\rm e}$ relation.

SOFTWARE CITATIONS

This work uses the following software packages:

(i) ASTROPY (Astropy Collaboration 2013; Price-Whelan et al. 2018)
(ii) COLOSSUS (Diemer 2018)
(iii) CORNERS.PY (Foreman-Mackey 2016)
(iv) DYNESTY (Speagle 2020)
(v) MATPLOTLIB (Hunter 2007)
(vi) NUMBA (Lam, Pitrou & Seibert 2015)
(vii) NUMPY (van der Walt, Colbert & Varoquaux 2011)
(viii) PYAUTOFIT (Nightingale, Hayes & Griffiths 2021b)
(ix) PYAUTOGALAXY (Nightingale et al. 2023)
(x) PYAUTOLENS (Nightingale & Dye 2015; Nightingale et al. 2018, 2021a)
(xi) PYQUAD (Kelly 2020)
(xii) PYTHON (Van Rossum & Drake 2009)
(xiii) SCIKIT-IMAGE (Van der Walt et al. 2014)
(xiv) SCIKIT-LEARN (Pedregosa et al. 2011)
(xv) SCIPY (Virtanen et al. 2020)
(xvi) SQLITE (Hipp 2020)

ACKNOWLEDGEMENTS

JWN is supported by the UK Space Agency, through grant ST/N001494/1, and a Royal Society Short Industry Fellowship. RJM is supported by a Royal Society University Research Fellowship and by the Science and Technologies Facilities Council (STFC) via grant ST/T002565/1, and the UK Space Agency via grant ST/W002612/1. JAK acknowledges support from a National Aeronautics and Space Administation (NASA) Postdoctoral Program Fellowship. AE is supported by STFC via grants ST/R504725/1 and ST/T506047/1. AA and QH acknowledge support from the European Research Council (ERC) through Advanced Investigator grant DMIDAS (GA

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

786910). This work used both the Cambridge Service for Data Driven Discovery (CSD3) and the Distributed Research using Advanced Computing (DiRAC) Data-Centric system, project code dp195, which are operated by the University of Cambridge and Durham University on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). These were funded by Business, Innovation and Skills (BIS) capital grant ST/K00042X/1, STFC capital grants ST/P002307/1, ST/R002452/1, ST/H008519/1, ST/K00087X/1, STFC Operations grants ST/K003267/1, ST/K003267/1, and Durham University. DiRAC is part of the UK National E-Infrastructure.

DATA AVAILABILITY

Text files and images of every model fit performed in this work are available at https://github.com/Jammy2211/autolens_abell_1201. Full `DYNESTY` chains of every fit are available at https://github.com/Jammy2211/autolens_abell_1201.

REFERENCES

Ade P. A. et al., 2016, A&A, 594, A13
Amorisco N. C. et al., 2022, MNRAS, 510, 2464
Astropy Collaboration, 2013, A&A, 558, A33
Banik U., van den Bosch F. C., Tremmel M., More A., Despali G., More S., Vegetti S., McKean J. P., 2019, MNRAS, 483, 1558
Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Gavazzi R., Moustakas L. A., Wayth R., Schlegel D. J., 2008, ApJ, 682, 964
Bolton A. S. et al., 2012, ApJ, 757, 82
Cao X. et al., 2021, Resea. Astron. Astrophys., 22, 30
Chen M. C., Broadhurst T., Lim J., Diego J. M., Ohyama Y., Ford H., Benítez N., 2018, ApJ, 863, 135
Collett T. E., 2015, ApJ, 811, 20
Davis T. A., Bureau M., Onishi K., Cappellari M., Iguchi S., Sarzi M., 2017, MNRAS, 468, 4675
Diemer B., 2018, ApJS, 239, 35
Dullo B. T., 2019, ApJ, 886, 80
Dullo B. T., Graham A. W., 2013, ApJ, 768, 36
Dullo B. T., Graham A. W., 2014, MNRAS, 444, 2700
Dye S., Evans N. W., Belokurov V., Warren S. J., Hewett P., 2008, MNRAS, 388, 384
Edge A. C., Smith G. P., Sand D. J., Treu T., Ebeling H., Allen S. W., van Dokkum P. G., 2003, ApJ, 599, L69
Etherington A. et al., 2022a, MNRAS, 517, 3275 (E22)
Etherington A. et al., 2022b, preprint (arXiv:2207.04070)
Event Horizon Telescope Collaboration, 2019, ApJ, 875, L1
Faber S. M. et al., 1996, AJ, 114, 1771
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Ferreras I. et al., 2019, MNRAS, 489, 608
Foreman-Mackey D., 2016, J. Open Source Softw., 1, 24
Gebhardt K. et al., 2000, ApJ, 539, L13
Graham A. W., 2001, AJ, 121, 820
Graham A. W., 2012, ApJ, 746, 113
He Q. et al., 2022a, MNRAS, 512, 5862
He Q. et al., 2022b, MNRAS, 511, 3046
He Q. et al., 2023, MNRAS, 518, 220
Heckman T. M., Best P. N., 2014, ARA&A, 52, 589
Hernquist L., 1990, ApJ, 356, 359
Hezaveh Y. D., Marshall P. J., Blandford R. D., 2015, ApJ, 799, L1
Hezaveh Y. D. et al., 2016, ApJ, 823, 37
Hipp R. D., 2020, SQLite, available at: http://citebay.com/how-to-cite/sqlite/
Hlavacek-Larrondo J., Fabian A. C., Edge A. C., Hogan M. T., 2012, MNRAS, 424, 224
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jackson N., Tagore A. S., Roberts C., Sluse D., Stacey H., Vives-Arias H., Wucknitz O., Volino F., 2015, MNRAS, 454, 287
Keeton C. R., 2003, ApJ, 582, 17
Kelly A. J., 2020, pyquad, Zenodo, available at: https://zenodo.org/record/3936959#.ZAU62nZBzIV
Kochanek C. S., 2004, preprint (arXiv:astro-ph/0407232)
Koopmans L. V. et al., 2009, ApJ, 703, L51
Kormendy J., Bender R., 2013, ApJ, 769, L5
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, ApJS, 182, 216
La Barbera F. et al., 2019, MNRAS, 489, 4090
Lam S. K., Pitrou A., Seibert S., 2015, Proc. 2nd Workshop on the LLVM Compiler Infrastructure in HPC – LLVM '15. ACM, p. 1
Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, MNRAS, 460, 1214
Ma C. J., Owers M., Nulsen P. E., McNamara B. R., Murray S. S., Couch W. J., 2012, ApJ, 752, 139
Mahler G., Natarajan P., Jauzac M., Richard J., 2023, MNRAS, 518, 54
Mao S., Witt H. J., Koopmans L. V., 2001, MNRAS, 323, 301
McConnell N. J., Ma C. P., 2013, ApJ, 764, 184
McLure R. J., Dunlop J. S., 2004, MNRAS, 352, 1390
Merritt D., 2006, ApJ, 648, 976
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Nightingale J. W., Dye S., 2015, MNRAS, 452, 2940
Nightingale J. W., Dye S., Massey R. J., 2018, MNRAS, 478, 4738
Nightingale J. W., Massey R. J., Harvey D. R., Cooper A. P., Etherington A., Tam S. I., Hayes R. G., 2019, MNRAS, 489, 2049
Nightingale J. et al., 2021a, J. Open Source Softw., 6, 2825
Nightingale J. W., Hayes R. G., Griffiths M., 2021b, J. Open Source Softw., 6, 2550
Nightingale J. W. et al., 2022, preprint (arXiv:2209.10566)
Nightingale J. W. et al., 2023, J. Open Source Softw., 8, 4475
O'Riordan C. M., Warren S. J., Mortlock D. J., 2019, MNRAS, 487, 5143
O'Riordan C. M., Warren S. J., Mortlock D. J., 2020, MNRAS, 496, 3424
O'Riordan C. M., Warren S. J., Mortlock D. J., 2021, MNRAS, 501, 3687
Oguri M., 2021, PASP, 133, 074504
Oldham L. J., Auger M. W., 2018, MNRAS, 476, 133
Orban De Xivry G., Marshall P., 2009, MNRAS, 399, 2
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Peterson B. M. et al., 2004, ApJ, 613, 682
Postman M. et al., 2012, ApJ, 756, 159
Price-Whelan A. M. et al., 2018, AJ, 156, 123
Quinn J. et al., 2016, MNRAS, 459, 2394
Rines K., Geller M. J., Diaferio A., Kurtz M. J., 2013, ApJ, 767, 15
Rusin D., Ma C.-P., 2001, ApJ, 549, L33
Rusin D., Keeton C. R., Winn J. N., 2005, ApJ, 627, L93
Schneider P., 2019, preprint (arXiv:1409.0015)
Sersic J. L., 1968, Cordoba. Observatorio Astronomico, Argentina
Shen Y., 2013, Bull. Astron. Soc. India, 41, 61
Shu Y. et al., 2016, ApJ, 824, 86
Smith A., Bromm V., 2019, Contemp. Phys., 60, 111
Smith R. J., Lucey J. R., Edge A. C., 2017a, MNRAS, 467, 836
Smith R. J., Lucey J. R., Edge A. C., 2017b, MNRAS, 471, 383
Sonnenfeld A., Gavazzi R., Suyu S. H., Treu T., Marshall P. J., 2013, ApJ, 777, 97
Speagle J. S., 2020, MNRAS, 493, 3132
Suyu S. H., Marshall P. J., Hobson M. P., Blandford R. D., 2006, MNRAS, 371, 983
Tamura Y., Oguri M., Iono D., Hatsukade B., Matsuda Y., Hayashi M., 2015, PASJ, 67, 1
Tessore N., Benton Metcalf R., 2015, A&A, 580, A79
Tessore N., Bellagamba F., Metcalf R. B., 2016, MNRAS, 463, 3115
Thater S., Krajnović D., Nguyen D. D., Iguchi S., Weilbacher P. M., 2019, Proc. Int. Astron. Un., 14, 199
Thomas J., Saglia R. P., Bender R., Erwin P., Fabricius M., 2014, ApJ, 782, 39
Thomas J., Ma C. P., McConnell N. J., Greene J. E., Blakeslee J. P., Janish R., 2016, Nature, 532, 340
Trujillo I., Erwin P., Ramos A. A., Graham A. W., 2004, AJ, 127, 1917
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput Sci. Eng, 13, 22

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

van Dokkum P., Conroy C., Villaume A., Brodie J., Romanowsky A. J., 2017, ApJ, 841, 68
van den Bosch R. C. E., 2016, ApJ, 831, 134
Van der Walt S., Schönberger J. L., Nunez-Iglesias J., Boulogne F., Warner J. D., Yager N., Gouillart E., Yu T., 2014, PeerJ, 2, e453
Van Rossum G., Drake F. L., 2009, Python 3 Reference Manual. CreateSpace, Scotts Valley, CA
Vegetti S., Koopmans L. V., Bolton A., Treu T., Gavazzi R., 2010, MNRAS, 408, 1969
Vegetti S., Lagattuta D. J., McKean J. P., Auger M. W., Fassnacht C. D., Koopmans L. V., 2012, Nature, 481, 341
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Warren S., Dye S., 2003, ApJ, 590, 673 (WD03)
Winn J. H., Rusin D., Kochanek C. S., 2004, Nature, 427, 613
Wong K. C., Ishida T., Tamura Y., Suyu S. H., Oguri M., Matsushita S., 2017, ApJ, 843, L35
Zhao H., 1996, MNRAS, 278, 488

APPENDIX A: LENS PROFILES

A1 Priors

Non-linear search chaining (see Section 3.8) updates the priors on the lens model parameters throughout the SLaM pipelines. Table A1 lists the priors assumed for every mass model parameter in the Mass pipeline which our Bayesian evidences are based. Details of the specific prior used for every lens model parameter in every model fit are provided at https://github.com/Jammy2211/autolens_abell_1201, where the full sets of `DYNESTY` results are also provided.

The mass-to-light ratio of each stellar light model assumes log uniform priors, where the lower and upper limits correspond to values that give Einstein masses of 0.01 and 5 times the Einstein mass inferred for the total mass profile fitted previously. Radial gradients assume uniform priors between −0.2 and 1.0. The NFW dark matter

Table A1. The priors on every parameter for the mass profiles used in this work, when they are fitted in the Mass pipeline and therefore from which our final parameter estimates and Bayesian evidences are based. Column 1 gives the model name. Column 2 gives the parameter. Column 3 gives the type of prior. Column 4 gives the values of that prior for fits to the F390W image and column 5 to the F814W image. For uniform and log uniform priors l and u give the lower and upper limits assumed. For Gaussian priors μ is the centre of the Gaussian and σ its width.

Model	Parameter	Prior	Units	Values (F390W)	Values (F814W)
Elliptical	$x^{\rm mass}$	Gaussian	arcsec	$\mu = 0.036, \sigma = 0.05$	$\mu = 0.047, \sigma = 0.05$
Power law (PL)	$y^{\rm mass}$	Gaussian	arcsec	$\mu = 0.115, \sigma = 0.05$	$\mu = 0.070, \sigma = 0.05$
	$\epsilon_1^{\rm mass}$	Gaussian		$\mu = 0.094, \sigma = 0.2$	$\mu = 0.201, \sigma = 0.2$
	$\epsilon_1^{\rm mass}$	Gaussian		$\mu = 0.048, \sigma = 0.2$	$\mu = -0.016, \sigma = 0.2$
	$\theta_{\rm E}^{\rm mass}$	Gaussian	arcsec	$\mu = 1.964, \sigma = 0.491$	$\mu = 2.002, \sigma = 0.05$
	$\gamma^{\rm mass}$	Uniform		$l = 1.5, u = 3.0$	$l = 1.5, u = 3.0$
Broken	$x^{\rm mass}$	Gaussian	arcsec	$\mu = 0.036, \sigma = 0.05$	$\mu = 0.047, \sigma = 0.05$
power law (BPL)	$y^{\rm mass}$	Gaussian	arcsec	$\mu = 0.115, \sigma = 0.05$	$\mu = 0.070, \sigma = 0.05$
	$\epsilon_1^{\rm mass}$	Gaussian		$\mu = 0.094, \sigma = 0.2$	$\mu = 0.201, \sigma = 0.05$
	$\epsilon_1^{\rm mass}$	Gaussian		$\mu = 0.048, \sigma = 0.2$	$\mu = -0.016, \sigma = 0.05$
	$\theta_{\rm E}^{\rm mass}$	Gaussian	arcsec	$\mu = 1.964, \sigma = 0.491$	$\mu = 2.002, \sigma = 0.05$
	$t_1^{\rm mass}$	Uniform		$l = 0.3, u = 1.5$	$l = 0.3, u = 1.5$
	$t_2^{\rm mass}$	Uniform		$l = 0.3, u = 2.0$	$l = 0.3, u = 2.0$
	$r_{\rm B}^{\rm mass}$	Uniform	arcsec	$l = 0.0, u = 1.0$	$l = 0.0, u = 1.0$
Sersic x3	$\Psi^{\rm bulge}$	Log Uniform	$e^- s^{-1}$	$l = 0.200, u = 100.163$	$l = 0.237, u = 118.31$
	$\Psi^{\rm disc}$	Log Uniform	$e^- s^{-1}$	$l = 0.048, u = 24.500$	$l = 0.105, u = 52.57$
	$\Psi^{\rm env}$	Log Uniform	$e^- s^{-1}$	$l = 0.092, u = 46.200$	$l = 0.315, u = 157.66$
Sersic x2	$\Psi^{\rm bulge}$	Log Uniform	$e^- s^{-1}$	$l = 0.200, u = 98.872$	$l = 0.233, u = 116.60$
	$\Psi^{\rm disc}$	Log Uniform	$e^- s^{-1}$	$l = 0.035, u = 17.704$	$l = 0.079, u = 39.68$
All sersics	Γ	Uniform		$l = -0.2, u = 1.0$	$l = -0.2, u = 1.0$
Elliptical NFW	$x^{\rm dark}$	Gaussian	arcsec	$\mu = 0.0, \sigma = 0.1$	$\mu = 0.0, \sigma = 0.1$
	$y^{\rm dark}$	Gaussian	arcsec	$\mu = 0.0, \sigma = 0.1$	$\mu = 0.0, \sigma = 0.1$
	$\epsilon_1^{\rm dark}$	Gaussian		$\mu = 0.0, \sigma = 0.2$	$\mu = 0.0, \sigma = 0.2$
	$\epsilon_1^{\rm dark}$	Gaussian		$\mu = 0.0, \sigma = 0.2$	$\mu = 0.0, \sigma = 0.2$
	$M_{200}^{\rm dark}$	Log Uniform	$M_\odot$	$l = 1 \times 10^{10}, u = 1 \times 10^{15}$	$l = 1 \times 10^{10}, u = 1 \times 10^{15}$
Point mass	$\theta_{\rm E}^{\rm smbh}$	Uniform	arcsec	$l = 0.0\ u = 3.0$	$l = 0.0\ u = 3.0$
shear	$\epsilon_1^{\rm ext}$	Gaussian		$\mu = -0.141, \sigma = 0.05$	$\mu = -0.105, \sigma = 0.05$
	$\epsilon_2^{\rm ext}$	Gaussian		$\mu = 0.236, \sigma = 0.05$	$\mu = 0.210, \sigma = 0.05$

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

profile is parametrized with its normalization as the mass at two hundred times the critical density of the Universe, $M_{200}^{\rm dark}$, and assumes a log uniform prior between 10^9 $M_\odot$ and 10^{15} $M_\odot$.

Identical prior passing is used in the Source and Light pipelines as in Etherington et al. (2022a) and we also use the likelihood cap described in this work to infer errors on lens model parameters, with all errors quoted at a 3σ confidence interval unless stated otherwise.

Due to prior passing, the prior on the BPL centre is not (0.0 arcsec, 0.0 arcsec) but offset to (0.036 arcsec, 0.115 arcsec) for the F390W fit. An important aspect of our results is that we infer a BPL centre that is offset from the luminous bulge, which we argue is non-physical. We verify that manually setting this prior to be centred on (0.0 arcsec, 0.0 arcsec) does not infer an accurate model that is not offset from the bulge light (these results are included at https://github.com/Jammy2211/autolens_abell_1201). In fact, these fits infer much lower Bayesian evidences.

APPENDIX B: LIGHT MODELS

This section presents the results of fitting the F390W and F814W images of Abell 1201 with different light models, which is performed in the Light pipeline. Fig. B1 shows projected 2D images of each model light profile, for the single Sersic model (top left-hand panel), the double Sersic model assuming no geometric alignments (top-centre and top-right panels) and the triple Sersic model (bottom panels). The single Sersic model infers a compact central bulge

Table B1. The Bayesian Evidence, ln $\mathcal{Z}$, of each model fit performed by the Light pipeline, which compares models with one, two, or three Sersic profiles. Fits to both the F814W and F390W images are shown. Models that make different assumptions for the alignment of the (x, y) centre and (ϵ_1, ϵ_2) elliptical components of the bulge, disc, and envelope are shown. A tick mark indicates that this assumption is used in the model, for example the second row is a model where both the elliptical components and centres are aligned. The triple Sersic model for the F390W is omitted because it went to unphysical solutions where one Sersic component was used to fit structure in the lensed source.

Filter	Number of Sersics	Aligned elliptical components	Aligned centres	Evidence
F814W	1	N/A	N/A	76 664.15
F814W	2	✓	✓	77 616.09
F814W	2	✗	✓	78 049.53
F814W	2	✗	✗	78 181.48
F814W	3	✗	✗	**78 193.40**
F390W	1	N/A	N/A	123 604.22
F390W	2	✓	✓	123 962.79
F390W	2	✗	✓	124 275.33
F390W	2	✗	✗	**124 663.78**
F390W	3	✗	✗	N/A

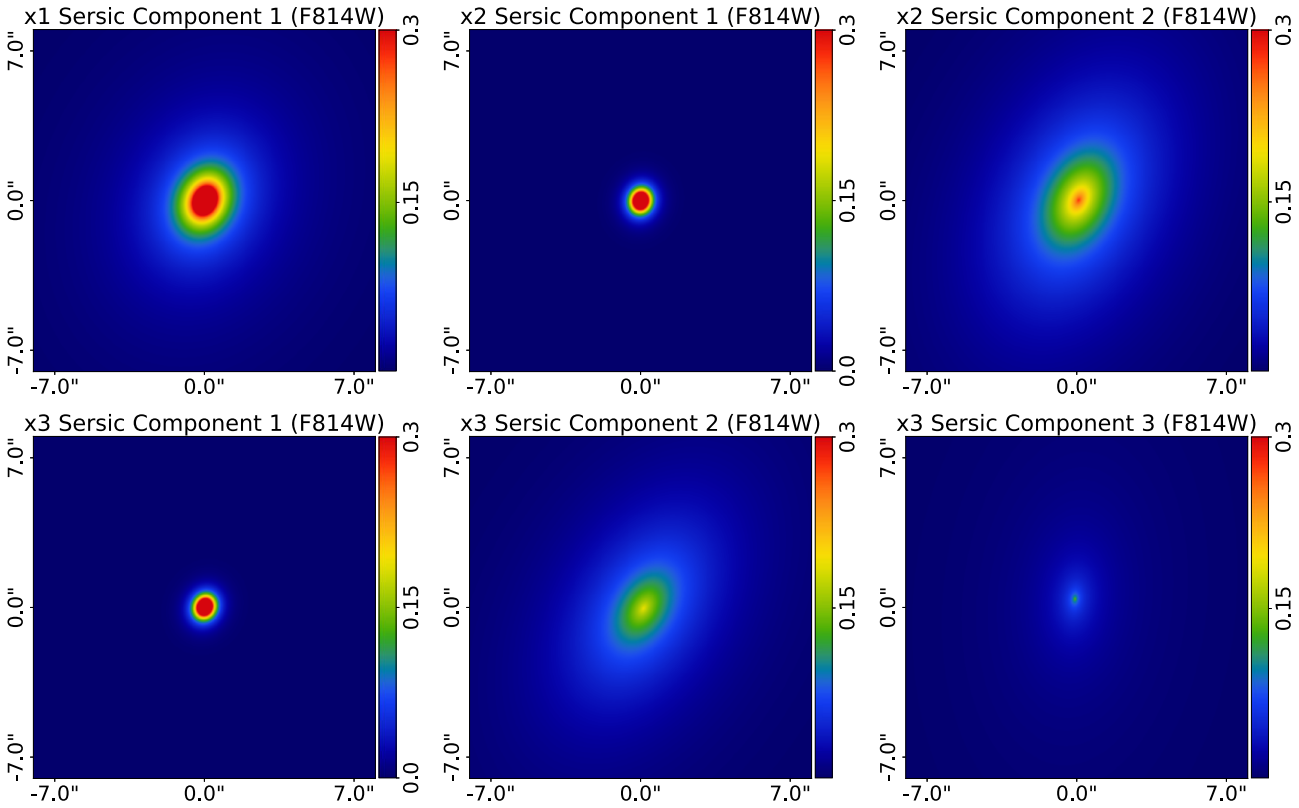

Figure B1. 2D projections of the individual light profiles for the following three lens galaxy light models: (i) a single Sersic profile (top left-hand panel); (ii) a double Sersic profile where the centres and elliptical components are not aligned, representing a central bulge (top middle panel), and extended component (top right-hand panel); (iii) a triple Sersic model where no geometric components are aligned, representing a bulge, an extended component and a third inner component (bottom row). All models are fitted with a fixed isothermal mass profile with external shear and a pixelized source reconstruction which changes for every light profile fitted. Each intensity plot corresponds to the maximum likelihood light model of a model fit using the F814W image (the F390W image's blue wavelength makes it is less suited to tracing the lens galaxy's stellar mass).

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

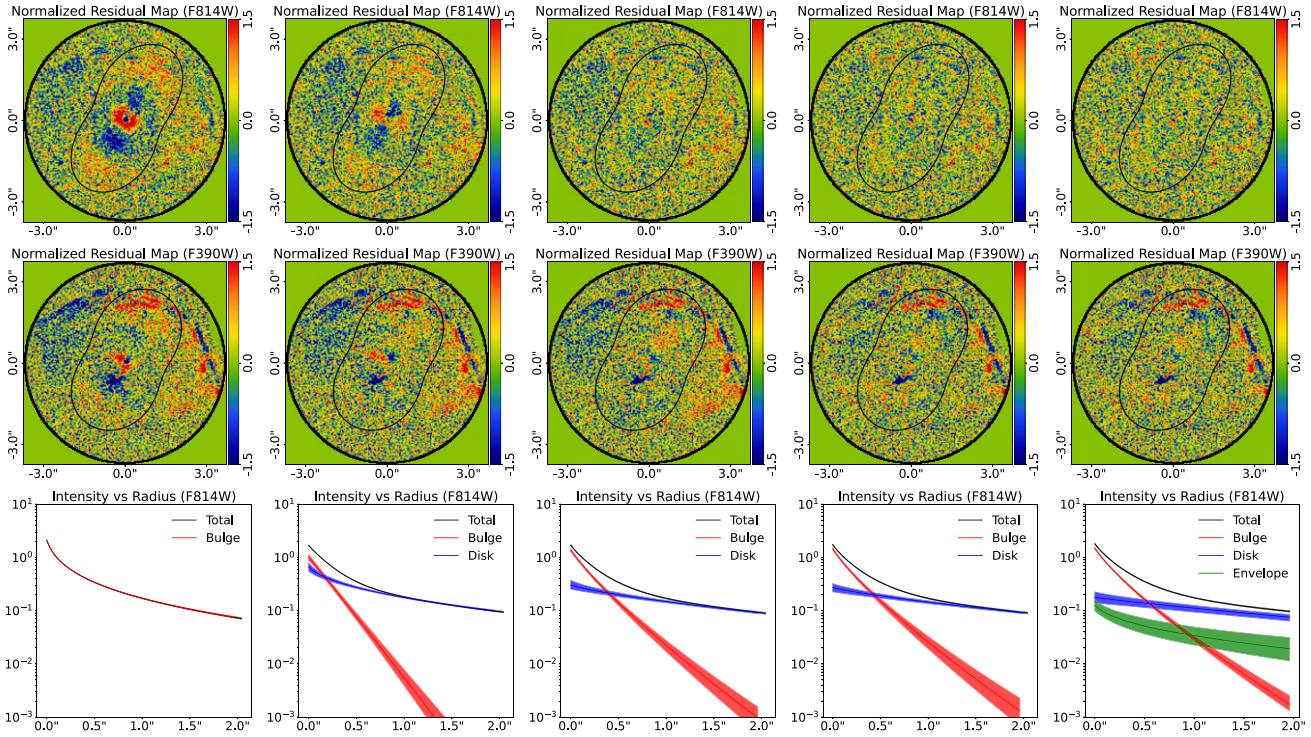

Figure B2. The normalized residuals of fits to the F814W image (top row), F390W image (middle row) and 1D decomposed intensity profiles (bottom row) of the following five lens galaxy light models model fits (from left to right): (i) a single Sersic profile; (ii) a double Sersic profile where the centres and elliptical components are aligned; (iii) where their centres are aligns but elliptical components are not; (iv) where neither components are aligned and; (v) a triple Sersic model where no geometric components are aligned. All models are fitted with a fixed isothermal mass profile with external shear and a pixelized source reconstruction which changes for every light profile fitted. 1D profiles are computed using coordinates that extend radially outwards from the centre of the light profile and are aligned with its major axis. Each plot corresponds to the maximum likelihood light model of a model fit. The shaded regions show estimates of each light profile within 3σ confidence intervals. The black line shows the tangential critical curve of the mass model, the white line the radial critical curve, and the black cross(es) towards the centre of each figure the centre(s) of each light profile component.

with Sersic index $n^{\rm bulge} \simeq 4$, consistent with a massive elliptical galaxy. The double Sersic model decomposes the lens galaxy's light into two distinct components, consisting of a compact bulge similar to the single Sersic fit but with a much lower value of $n^{\rm bulge} \simeq 1.25$, surrounded by a more extended and elliptical component where $n^{\rm disc} \simeq 1.3$. The half-light radius of this extended component is $R_{\rm eff}^{\rm disk} \simeq 5.0''$, well beyond the strong lensing features. The triple Sersic model infers these two components, but includes a fainter additional inner structure.

The Bayesian evidence values, $\ln \mathcal{Z}$, of the light models informs us which provides the best fit to the data. These are given for both F390W and F814W images in Table B1. Models assuming a single Sersic profile give significantly worse fits than those using multiple profiles, indicating it does not capture the extended component. Three models assuming two Sersic profiles are compared, where (i) their centre and elliptical components are aligned; (ii) their centres are aligned but elliptical components are not; and (iii) their centres are also free to vary. For both images, model (iii) is preferred, with a value of $\Delta \ln \mathcal{Z} > 100$ the other models for the F814W data. For the F814W image, a triple Sersic (with all geometric parameters free to vary) gives a value $\ln \mathcal{Z} = 11.55$ above that of the two Sersic model, indicating that it is the marginally favoured model.

Fig. B2 shows the normalized residuals of these fits. For the single Sersic model and models with geometric alignments residuals are evident around the lens galaxy's centre in both the F814W and F390W bands, consistent with the Bayesian evidences. In the F814W image, the double Sersic model with free centres and the triple Sersic model gave a significant increase in $\ln \mathcal{Z}$. However, the improvements are not visible in the residuals, indicating they improve the light model fractionally over many pixels.

The lower panels of this figure show 1D plots of the intensity as a function of radius for each component. The inner structure contributes to most of the stellar light within ∼1.0 arcsec where the counter image is observed, whereas at the location of the giant arc the extended component makes up over 95 per cent of the total emission. They also show that the outer component makes up the majority of the lens galaxy's total luminous emission, albeit most is beyond the 3.0 arcsec radius where the lensed source is constrained.

APPENDIX C: DOUBLE SERSIC MODELS

The results of fitting the decomposed model with two Sersic profiles are shown in Figs C1 and C2. These figures follow the same layout as Figs 6 and 7 in the main paper. Results show the same behaviour as the triple Sersic fitted in the main paper, including extraneous flux in the counter image reconstruction when the model omits an SMBH. The double Sersic fit with an SMBH has a $\ln \mathcal{Z}$ value 29.93 below the triple Sersic with an SMBH. This suggests that the lensing effects of the faint inner structure the third Sersic represents plays a role in reconstructing the counter image.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

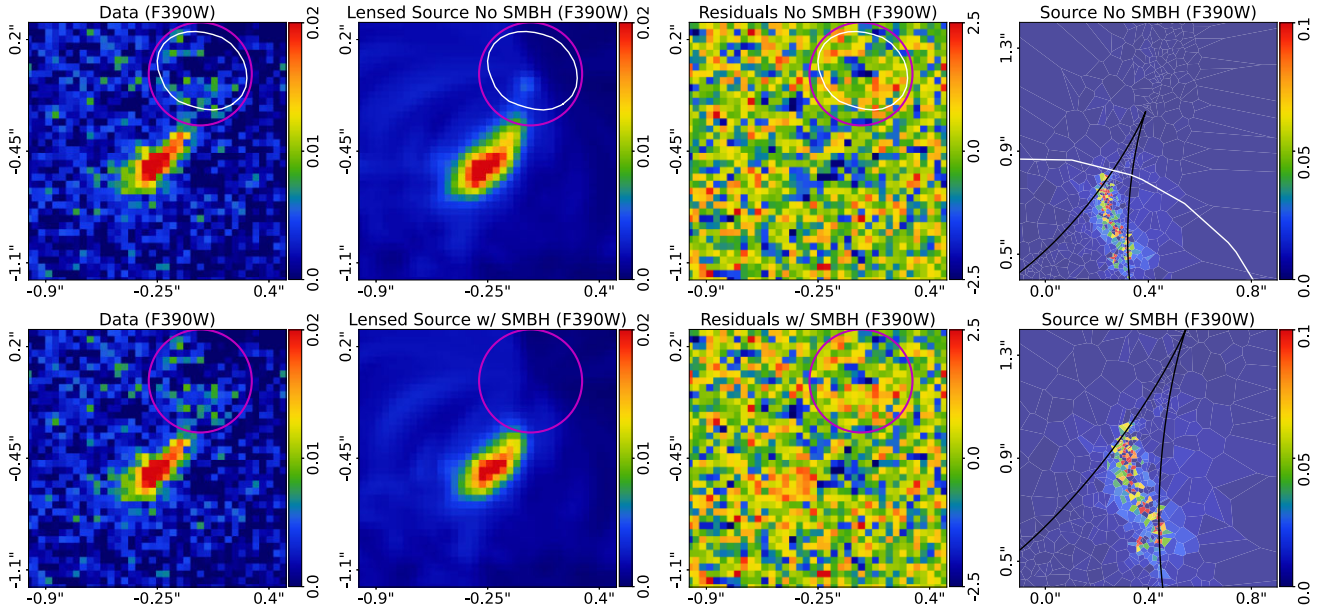

Figure C1. Zoom-ins of the observed counter image in the F390W data (left-hand panel), the model lensed source (left-centre panel), the normalized residuals (right-centre panel) and the source reconstruction (right panel). The top and bottom rows shows double Sersic plus NFW decomposed model fits without and with an SMBH respectively. All models include an external shear. Models that omit an SMBH form an additional clump of light in the counter image, which is not present in the data. The tangential caustic is shown by a black line and radial critical curve and caustic a white line; the latter does not form for models including an SMBH.

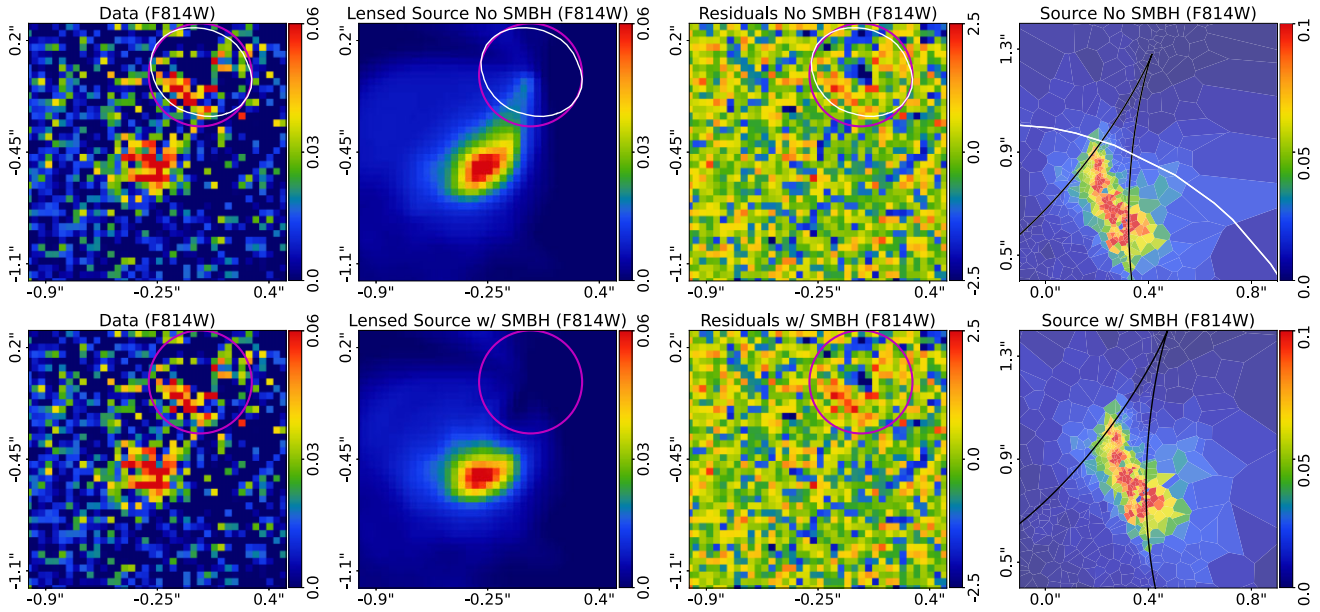

Figure C2. The same as Fig. C1 but for the F814W data.

APPENDIX D: TOTAL MASS MODELS

This appendix shows the results of fitting two total mass models: the PL (Tessore & Benton Metcalf 2015) and BPL (O'Riordan et al. 2019, 2020, 2021). Like in the main paper, we compare fits with and without a point-mass representing an SMBH. We focus on the Bayesian evidence, $\ln \mathcal{Z}$, and the reconstruction of the counter image. We investigate whether the extraneous flux removed by the SMBH for the decomposed models can be removed by either of these profiles without an SMBH. The inferred model parameters for the PL and BPL models are given in Tables D2 and D3.

D1 Profile equations

A softened PL ellipsoid density profile of form

$$\kappa^{\rm mass}(\xi) = \frac{3-\gamma^{\rm mass}}{1+q^{\rm mass}} \left(\frac{\theta_{\rm E}^{\rm mass}}{\xi} \right)^{\gamma^{\rm mass}-1} \quad \text{(D1)}$$

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

Table D1. The Bayesian Evidence, $\ln \mathcal{Z}$, of each model fit using total mass models that collectively represent the lens's stellar and dark matter, where all models also include an external shear. $\ln \mathcal{Z}$ values for both the F390W and F814W images are shown. The favoured models given our criteria of $\Delta \ln \mathcal{Z} > 10$ is shown in bold. The PL mass model without an SMBH produces lower values of $\ln \mathcal{Z}$ than the PL model with an SMBH and both BPL models. The BPL model without an SMBH produces a $\ln \mathcal{Z}$ comparable to all models including an SMBH. The PL models favoured by model comparison are shown in bold; no bold model is shown for the BPL because models with and without an SMBH are both within the threshold of $\Delta \ln \mathcal{Z} > 10$ of one another.

Filter	Model	Includes? SMBH	$\ln \mathcal{Z}$
F390W	PL	✗	125 562.45
F390W	PL	✓	**125 707.20**
F390W	BPL	✗	125 699.90
F390W	BPL	✓	125 693.78
F814W	PL	✗	78 301.58
F814W	PL	✓	**78 330.39**
F814W	BPL	✗	78 331.17
F814W	BPL	✓	78 329.28

is assumed, where $\theta_{\rm E}^{\rm mass}$ is the model Einstein radius in arcseconds. The PL density slope is $\gamma^{\rm mass}$, and setting $\gamma^{\rm mass} = 2$ gives the singular isothermal ellipsoid (SIE) model. Deflection angles for the PL are computed via an implemention of the method of (Tessore & Benton Metcalf 2015) in `PYAUTOLENS`.

We also use the elliptical BPL profile (O'Riordan et al. 2019, 2020, 2021) with convergence

$$\kappa^{\rm mass}(r) = \begin{cases} \theta_{\rm E}^{\rm mass} \left(r_{\rm b}^{\rm mass}/r\right)^{t_1^{\rm mass}}, & r \leq r_{\rm b}^{\rm mass} \\ \theta_{\rm E}^{\rm mass} \left(r_{\rm b}^{\rm mass}/r\right)^{t_2^{\rm mass}}, & r > r_{\rm b}^{\rm mass} \end{cases}, \tag{D2}$$

where $r_{\rm b}^{\rm mass}$ is the break radius, $\theta_{\rm E}^{\rm mass}$ is the convergence at the break radius, $t_1^{\rm mass}$ is the inner slope, and $t_2^{\rm mass}$ is the outer slope. The isothermal case is given by $t_1^{\rm mass} = t_2^{\rm mass} = 1.0$.

D2 PL models

We first investigate fits using the simpler PL mass model. The PL parametrization has less flexibility in adjusting its central density compared to the BPL. The top two rows of Table D1 show the $\ln \mathcal{Z}$ values inferred for PL model fits with and without an SMBH. Models including an SMBH are strongly favoured, giving $\Delta \ln \mathcal{Z} = 145$ for the F390W data and $\Delta \ln \mathcal{Z} = 29$ for the F814W.

Fig. D1 shows zoom-ins of the PL model's reconstruction of the counter image. The figure shows the same behaviour seen for the decomposed model in the main paper, whereby the PL model without an SMBH produces central extraneous flux, which the inclusion of the SMBH removes. Fig. D2 shows this also occurs in the F814W image. The residuals of this extraneous flux are more significant than seen for the decomposed model fitted in the main paper, because of the PL model's reduced flexibility in adjusting its central density.

When the PL mass model includes an SMBH a value of $M_{\rm BH} = 3.83^{+1.56}_{-1.72} \times 10^{10}\,{\rm M}_\odot$ is inferred, which is consistent with the $M_{\rm BH}$ values inferred for the decomposed models. The SMBH changes the ray-tracing such that the lens model can now reproduce the counter image's structure accurately. The PL also infers a shallower slope of $\gamma^{\rm mass} = 1.65^{+0.12}_{-0.12}$, compared to the value $\gamma^{\rm mass} = 1.82^{+0.05}_{-0.05}$ inferred without an SMBH. The model without an SMBH therefore tries (and fails) to better fit the counter image by placing more mass centrally.

Fits using the PL therefore support the inclusion of an SMBH is the lens model, and their reconstruction of the counter image produces the same behaviour seen for the decomposed model in the main paper.

D3 BPL models

We now inspect fits using the BPL, which has much greater flexibility than the PL in controlling its inner density. The bottom two rows of table D1 show the $\ln \mathcal{Z}$ values inferred for BPL model fits with and without an SMBH. For the F390W image the $\ln \mathcal{Z}$ value for the BPL model without an SMBH is 125699.90; this is 6.22 above the BPL model with an SMBH. This value is also within $\Delta \ln \mathcal{Z} \sim 1$ of the decomposed models including an SMBH fitted in the main paper (see table 4).

Fig. D3 shows zoom-ins of the BPL model's reconstruction of the counter image. Irrespective of whether an SMBH is included in the model, the extraneous flux in the reconstructed counter image seen for decomposed models and the PL model without an SMBH is not produced. Fig. D4 shows this is also true for fits to the F814W image.

Fig. D5 shows the 1D convergence profiles for the BPL mass models with and without an SMBH. Shaded regions shows 3σ confidence intervals for each profile. The inner density (e.g. within 0.3 arcsec) of the BPL without an SMBH is steeper than the decomposed models fitted in the main paper (and also the PL models). The BPL is therefore able to remove extraneous flux from the the reconstructed counter image because it places more mass centrally than any other mass model. The BPL model including an SMBH infers a shallower density profile, because the SMBH performs the ray-tracing which fits the counter image.

Fits using the BPL model therefore raise the possibility that an SMBH is not required in the lens mass model.

D4 Decomposed model validation

The BPL fits show that if the mass model has a sufficiently high inner density then it can reconstruct the counter image accurately. We therefore check whether the decomposed models fitted in the main paper can place as much mass centrally as the BPL without requiring an SMBH. The blue dashed line in Fig. D5 shows that if the bulge of the triple Sersic model assumes a radial gradient parameter with the value $\Gamma^{\rm bulge} = 0.9$, its central density matches that of the BPL. The decomposed model parametrization therefore includes models with inner densities comparable to the BPL. We did not infer them because they correspond to lower likelihood solutions (our inferred value is $\Gamma^{\rm bulge} = 0.52^{+0.21}_{-0.32}$ at 3σ confidence). We verify this by fitting decomposed models where a uniform prior on $\Gamma^{\rm bulge}$ for the bulge component is placed between 0.85 and 0.95. The $\ln \mathcal{Z}$ values of this model with three and two Sersics are 125560.70 and 125630.27, respectively, well below the value of 125699.06 found for the triple Sersic decomposed model including an SMBH.

We also investigate models that make the central dark matter density comparable to that of the BPL. The green dashed line in Fig. D5 shows that an NFW profile with a concentration that is a 3.5σ positive outlier on the mass–concentration relation (Ludlow et al. 2016) has a central density close to the BPL. We therefore fit a triple Sersic decomposed models that includes the scatter from the mass–concentration relation $\sigma^{\rm dark}$ as a free parameter with a uniform prior between 2.5 and 4.0. We infer $\ln \mathcal{Z} = 125329.01$, significantly below nearly all model fits, with or without an SMBH.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

Table D2. The inferred geometric model parameters of the PL and BPL total mass models fitted to the F390W image in the Mass pipeline. Errors are given at 3σ confidence intervals.

Model	$x^{\rm mass}$ (arcsec)	$y^{\rm mass}$ (arcsec)	$\epsilon_1^{\rm mass}$	$\epsilon_2^{\rm mass}$	$\epsilon_1^{\rm ext}$	$\epsilon_2^{\rm ext}$
PL	$0.014^{+0.035}_{-0.041}$	$-0.054^{+0.053}_{-0.053}$	$0.107^{+0.030}_{-0.028}$	$-0.088^{+0.024}_{-0.003}$	$-0.113^{+0.019}_{-0.018}$	$0.152^{+0.018}_{-0.025}$
BPL	$0.045^{+0.031}_{-0.033}$	$0.063^{+0.065}_{-0.061}$	$-0.105^{+0.026}_{-0.029}$	$0.145^{+0.027}_{-0.032}$	$-0.105^{+0.039}_{-0.024}$	$0.145^{+0.025}_{-0.038}$
PL + SMBH	$0.050^{+0.035}_{-0.040}$	$0.063^{+0.045}_{-0.071}$	$0.104^{+0.028}_{-0.026}$	$-0.089^{+0.026}_{-0.037}$	$-0.112^{+0.026}_{-0.021}$	$0.147^{+0.019}_{-0.029}$
BPL + SMBH	$0.052^{+0.001}_{-0.001}$	$0.079^{+0.004}_{-0.009}$	$0.101^{+0.002}_{-0.001}$	$-0.083^{+0.003}_{-0.005}$	$-0.117^{+0.002}_{-0.001}$	$0.146^{+0.022}_{-0.041}$

Table D3. The inferred model parameters of the PL and BPL total mass models fitted to the F390W image in the Mass pipeline. Errors are given at 3σ confidence intervals.

Model	$\theta_{\rm Ein}^{\rm mass}$ (arcsec)	$\gamma^{\rm mass}$	$t_1^{\rm mass}$	$t_2^{\rm mass}$	$\theta_{\rm B}^{\rm mass}$ (arcsec)	$\theta_{\rm Ein}^{\rm smbh}$ (arcsec)
PL	$1.925^{+0.046}_{-0.040}$	$1.818^{+0.042}_{-0.077}$				
BPL	$1.869^{+0.039}_{-0.041}$		$1.13^{+0.34}_{-0.20}$	$0.64^{+0.09}_{-0.15}$	$0.45^{+0.09}_{-0.22}$	
PL + SMBH	$1.56^{+0.18}_{-0.19}$	$1.66^{+0.08}_{-0.10}$				$0.53^{+0.12}_{-0.14}$
BPL + SMBH	$1.6165^{+0.0017}_{-0.0092}$		$0.6920^{+0.0061}_{-0.0419}$	$0.6637^{+0.0009}_{-0.0036}$	$0.2096^{+0.0320}_{-0.0030}$	$0.5544^{+0.0107}_{-0.0013}$

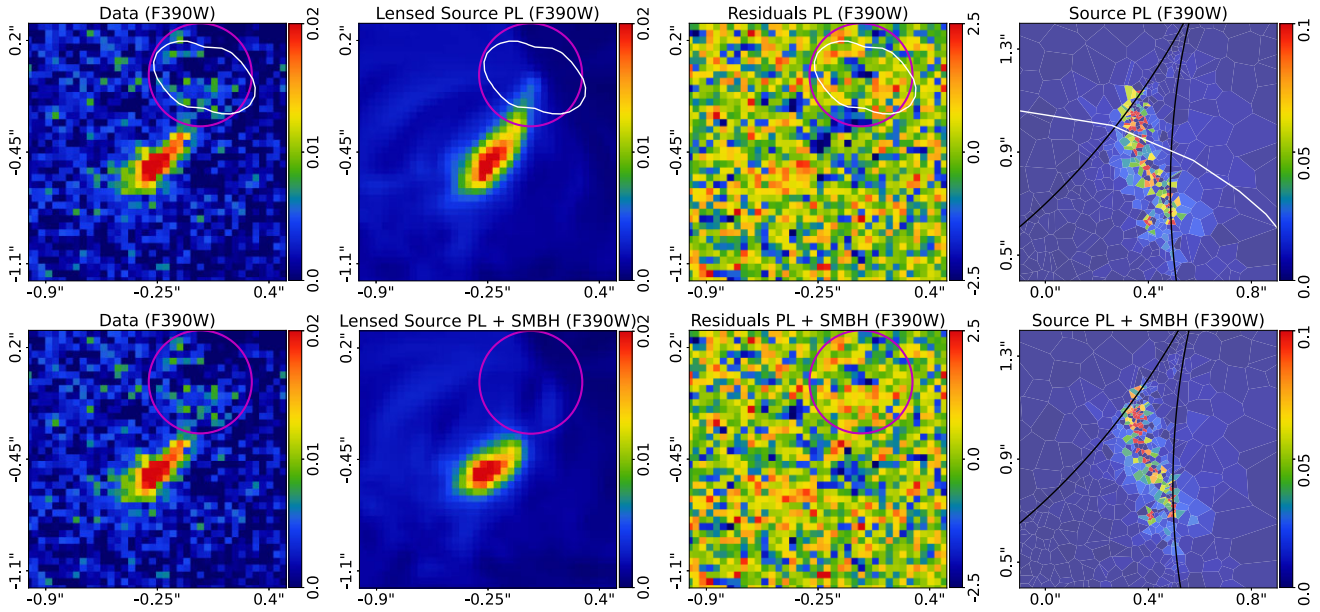

Figure D1. Zoom-ins of the observed counter image in the F390W data (left-hand panel), the model lensed source (left-centre panel), the normalized residuals (right-centre panel), and the source reconstruction (right-hand panel). The top and bottom rows show the PL mass model without and with an SMBH, respectively. All models include an external shear. Models that omit an SMBH form an additional clump of light in the counter image, which is not present in the data. The tangential caustic is shown by a black line and radial critical curve and caustic a white line; the latter does not form for models including an SMBH.

We therefore conclude that decomposed models that place as much mass centrally as the BPL model cannot attain a comparable $\ln \mathcal{Z}$ without an SMBH for fits to the F390W data. They are also unable to prevent extraneous flux appearing in the counter image.

D5 Mass model centring

The centre of the stellar mass component of the decomposed model is tied to that of the lens light, whereas the BPL has full freedom in choosing its centre. We now inspect the centring of the decomposed and BPL models in more detail, to see if any model appears more or less realistic or physically plausible. This will allow us to argue in favour or against the need for an SMBH.

Upon inspection of the different mass model parameters, fits using the BPL model (with or without an SMBH) infer mass model centres in the range $0.03 < x^{\rm mass} < 0.06$ and $0.04 < y^{\rm mass} < 0.09$ for the F390W image and $0.0 < x^{\rm mass} < 0.03$ and $0.02 < y^{\rm mass} < 0.07$ for the F814W image. Inspecting the lens light model fits, the inferred centre of the bulge at 3σ confidence is $x^{\rm bulge} = -0.008^{+0.003}_{-0.003}$ and $y^{\rm bulge} = 0.003^{+0.003}_{-0.003}$ for the F814W image and $x^{\rm bulge} = -0.013^{+0.007}_{-0.006}$ and $y^{\rm bulge} = 0.007^{+0.003}_{-0.003}$ for the F390W image. The BPL model is therefore shifting its centre $\geq 0.04''$ (a full pixel) away from the bulge centre, a shift which corresponds to ≥ 120 pc.

We now fit a BPL model without an SMBH where the centre is fixed to that of the bulge ($x^{\rm mass} = -0.008$ and $y^{\rm mass} = 0.003$). This model's fit to the F390W image infers $\ln \mathcal{Z} = 125317.26$,

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

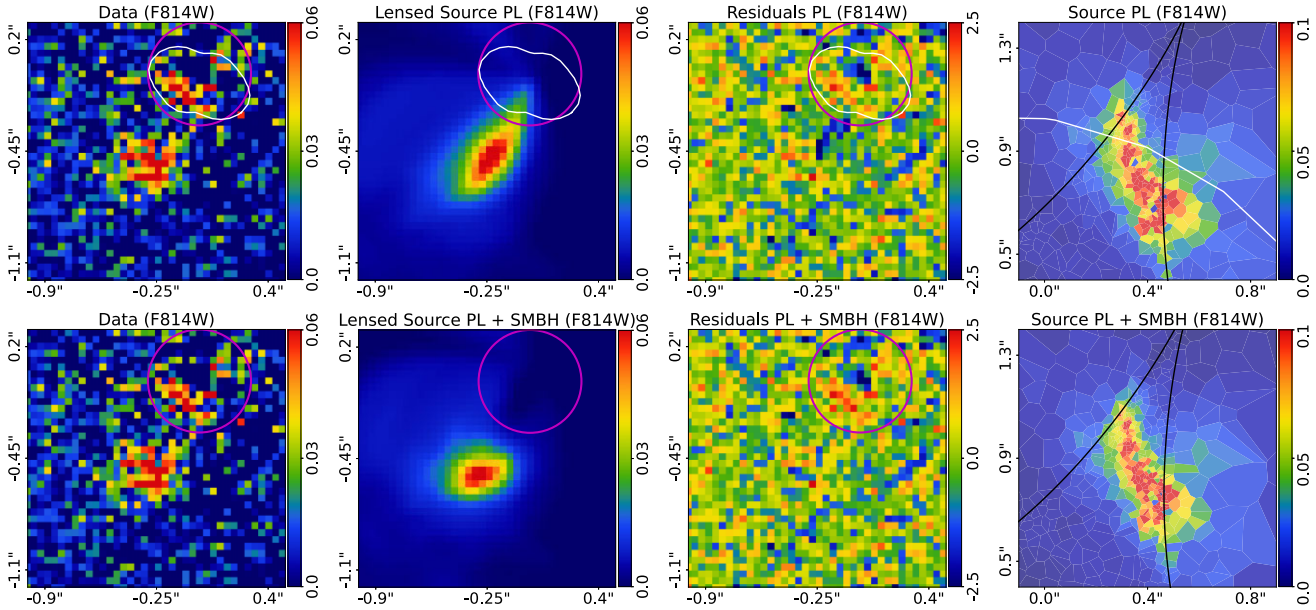

Figure D2. The same as Fig. D1 but for the F814W data.

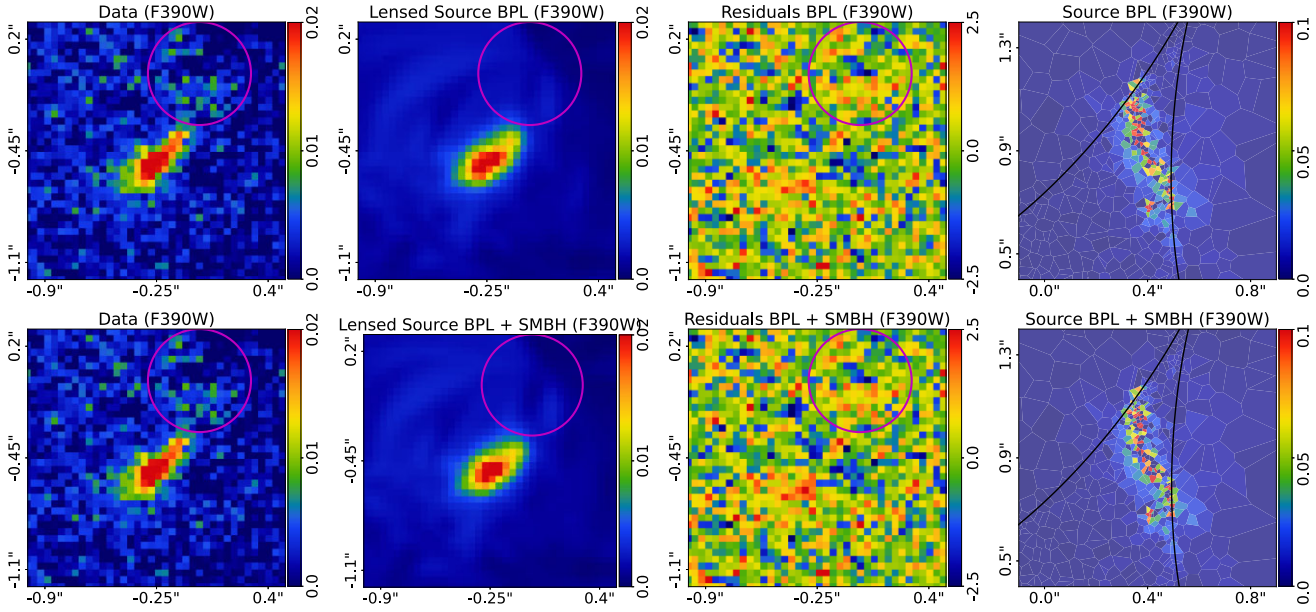

Figure D3. The same as Fig. D1 but for the BPL model and F390W data.

well below the value of $\ln \mathcal{Z} = 125699.90$ inferred for the BPL model with a free centre. When the BPL's centre is consistent with the luminous emission it therefore cannot reconstruct Abell 1201's source accurately.

We can now explain why decomposed models without an SMBH but with a bulge radial gradient around $\Gamma^{\rm bulge} = 0.9$ or a very concentrated dark matter halo did not give as high $\ln \mathcal{Z}$ values or remove extraneous flux from the reconstructed counter image. Even though their central density is as steep as the BPL model, steepening the mass profile only improves the overall fit when its centre is offset from the bulge by $\geq$120 pc in the positive x and y directions. Thus, not only does the BPL show a nonphysical offset from the bulge, but its ability to reconstruct the counter image accurately is dependent on the existence of this offset.

We therefore view the decomposed models with an SMBH fitted in the main paper as more reliable than the BPL model without an SMBH and discard the BPL model as non-physical.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

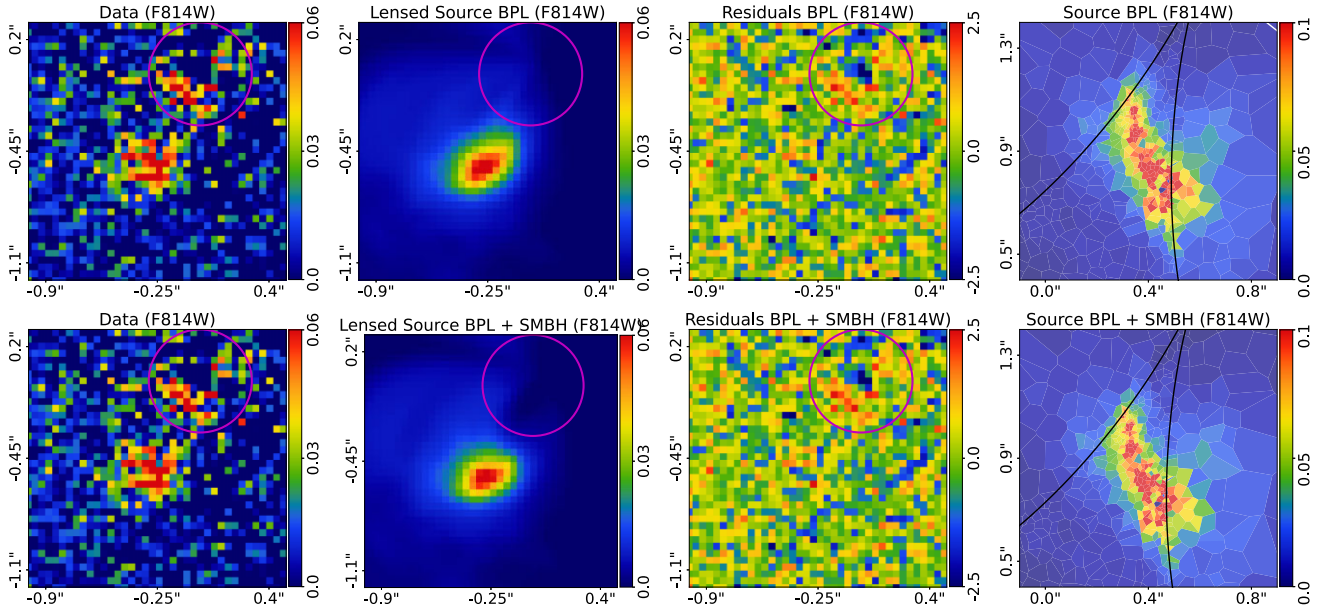

Figure D4. The same as Fig. D1 but for the BPL model and F814W data.

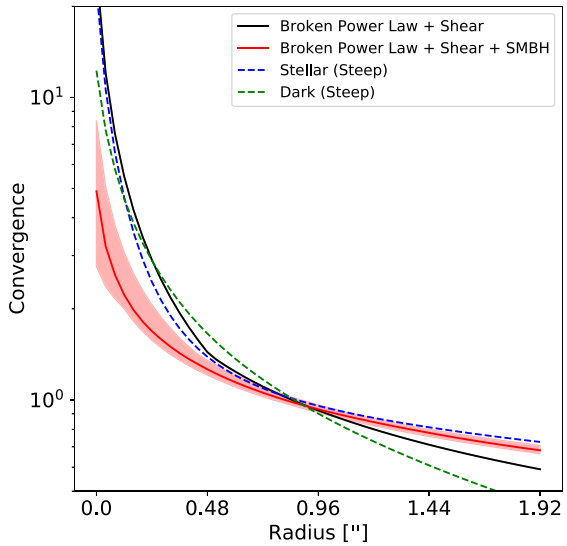

Figure D5. The convergence as a function of radius inferred using the F390W image for the total mass models: (i) the PL (black); (ii) the BPL (red); (iii) the PL and SMBH (blue) and; (iv) the BPL and SMBH (green), where all models include an external shear. Each line is computed using coordinates that extend radially outwards from the centre of the mass profile and are aligned with its major axis. Shaded regions for each mass model's convergence are shown, corresponding to the inferred 3σ confidence intervals. The BPL model places more mass centrally than all other models, consistent with its ability to reconstruct the counter image accurately.

APPENDIX E: LINE-OF-SIGHT GALAXY

Fig. 1 shows line-of-sight emission towards the right of the giant arc, around (4.0″, 1.0″). Smith et al. (2017a) show that this is a $z = 0.273$ galaxy, which is therefore located between the lens and

Table E1. The same as Table D1 but with the $z = 0.273$ galaxy around (4.0 arcsec, 1.0 arcsec) included in the lens galaxy mass model. This table shows fits assuming a decomposed mass model with two and three Sersic profiles.

Filter	Number of sersics	Includes SMBH?	ln $\mathcal{Z}$
F390W	2	✗	125 559.11
F390W	3	✓	125 608.05
F390W	2	✗	125 588.03
F390W	3	✓	125 596.66
F814W	2	✗	78 327.28
F814W	3	✓	78 322.25
F814W	2	✗	78 318.61
F814W	3	✓	78 316.00

Table E2. The same as Table D1 but with the $z = 0.273$ galaxy around (4.0 arcsec, 1.0 arcsec) included in the lens galaxy mass model. This table shows fits assuming a PL and BPL lens model.

Filter	Model	Includes SMBH?	ln $\mathcal{Z}$
F390W	PL	✗	125 434.63
F390W	PL	✓	125 589.40
F390W	BPL	✗	125 239.03
F390W	BPL	✓	125 586.22
F814W	PL	✗	78 264.73
F814W	PL	✓	78 323.10
F814W	BPL	✗	78 314.53
F814W	BPL	✓	78 311.49

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

source galaxies. The emission seen in the *HST* imaging appears as two (or more) distinct blobs. The [O III] emission shows similar structure indicating this is likely a single galaxy. We fit additional lens models to Abell 1201 that include this galaxy in the lens model as a spherical isothermal mass profile (see equation D1) where $\gamma^{\rm mass} = 2$), accounting for multiplane ray-tracing effects (Schneider 2019). The centre of this model is fixed to (3.6″, 0.95″) in the image-plane, which is updated when performing multiplane ray tracing. The $\ln \mathcal{Z}$ of these model fits are given in Tables E1 and E2. All models produce lower $\ln \mathcal{Z}$ values than those inferred in the main paper, indicating that including the galaxy does improve the lens model.

APPENDIX F: SMBH WITH FREE CENTRE

The $\ln \mathcal{Z}$ of model fits where the SMBH centre is free to vary are given in Tables F1 and F2. In agreement with the main paper's results, decomposed models including an SMBH with a free centre produce $\ln \mathcal{Z}$ increases at least 25 above decomposed models without an SMBH. For the triple Sersic decomposed model, the model whose SMBH centre is free gives $\ln \mathcal{Z} = 125665.66$ compared to $\ln \mathcal{Z} = 125699.06$ when the SMBH centre is fixed to the bulge centre. We interpret this decrease as a consequence of Occam's Razor (see Section 3.9), whereby the use of a too complex model that does not improve the fit to the data is being penalized. We find that the estimate of $M_{\rm BH}$ does not change when the SMBH centre is free to vary.

Table F1. The same as Table D1 but the SMBH centre is free to vary. This table shows fits assuming a decomposed mass model with two and three Sersic profiles.

Filter	Number of Sersics	Includes SMBH?	$\ln \mathcal{Z}$
F390W	2	✗	125 637.18
F390W	2	✓	125 665.66
F390W	3	✗	125 598.48
F390W	3	✓	125 661.03
F814W	2	✗	78 330.51
F814W	2	✓	78 327.26
F814W	3	✗	78 329.19
F814W	3	✓	78 324.19

Table F2. The same as Table D1 but the SMBH centre is free to vary. This table shows fits assuming a PL and BPL lens model.

Filter	Model	Includes SMBH?	$\ln \mathcal{Z}$
F390W	PL	✗	125 562.45
F390W	PL	✓	**125 683.77**
F390W	BPL	✗	125 699.90
F390W	BPL	✓	125 557.48
F814W	PL	✗	78 301.58
F814W	PL	✓	**78 321.82**
F814W	BPL	✗	78 331.17
F814W	BPL	✓	78 323.26

APPENDIX G: MODELS WITH SHALLOW INNER DENSITY

We encountered an alternative family of solutions that are characterized by (i) a shallow inner density profile that forms a larger radial critical curve than the solutions presented in the main paper, which cuts through the inner regions of the counter image; and (ii) the counter image reconstruction producing a pair of merging images (the models in the main paper reconstruct a single counter image). An example of such a model is shown in Fig. G1.

For decomposed models, these solutions are found when the radial gradient parameters (e.g. $\Gamma^{\rm bulge}$) are below zero and there is less mass relative to light. The low Sersic indices of the lens galaxy's light profiles ($n^{\rm bulge} = 1.28$ and $n^{\rm disc} = 1.16$) also help to produce a shallow inner density. For the BPL model, these correspond to solutions where the inner slope $t_1^{\rm mass} \sim 0.0$, the outer slope $t_2^{\rm mass} \sim 0.7$, and the break radius is $r_{\rm B}^{\rm mass} \sim 0.25''$. We verify that this family of models without an SMBH do not fit the data as well as models with an SMBH by performing `DYNESTY` fits, where the priors on certain mass-model parameters are constrained to uniform priors that restrict the analysis to these solutions. The priors can be found at https://github.com/Jammy2211/autolens_abell_1201.

The maximum likelihood solution for the double Sersic decomposed mass model are shown in Fig. G2. The reconstructed counter image is split in two, and fails to capture the appearance of the counter image in the data. For this model, the log Bayesian evidence value is $\mathcal{Z} =\sim$ 125 649, which is significantly below models with an SMBH which have a log evidence of $\mathcal{Z} =\sim$ 125 699. Table G1 compares the log Bayesian evidence values for the BPL model fits with a shallower inner density and also includes the values for the F814W. For both the F390W and F814W images, these solutions provide significantly worse fits to the data than models including an SMBH, confirming that they are ruled out by the data.

These fits also confirm that the central emission seen in the F814W data (Fig. 4; within magenta circle) is not a central image. Lens model fits using cored mass profiles would reconstruct the counter image, if it were the physically correct solution. The fact these solutions are not inferred confirms it is not a central image.

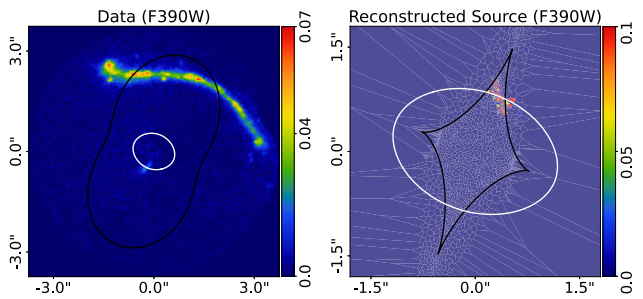

Figure G1. An example model where the lens mass model has a shallow inner density, which forms a larger radial critical curve than solutions presented in the main paper. The left-hand panel shows the observed data, with radial and tangential critical curves (white and black respectively) overlaid. The right-hand panel shows the corresponding source plane and source reconstruction, with the radial and tangential caustics (white and black respectively) overlaid.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023

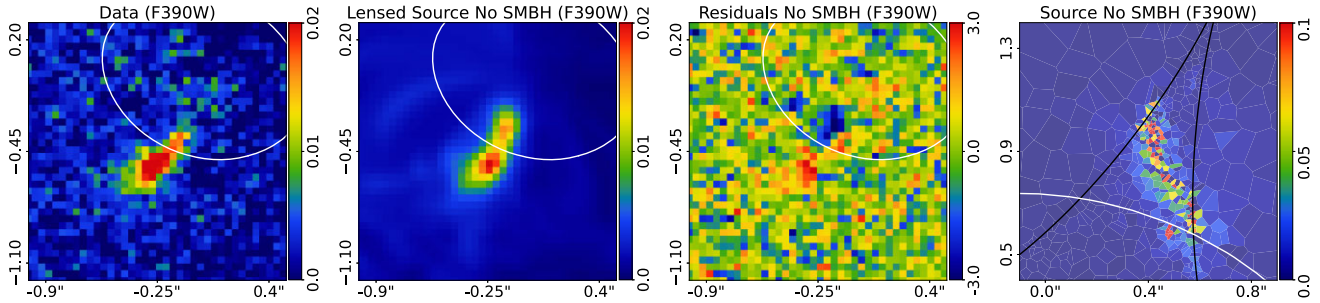

Figure G2. Zoom-ins of the observed counter image in the F390W data (left-hand panel), the model lensed source (left-centre panel), the normalized residuals (right-centre panel) and the source reconstruction (right-hand panel). These results are for the double Sersic plus NFW decomposed modelfits without an SMBH, where the parameter priors allow for solutions with a shallow inner density and large radial critical curve. The tangential caustic is shown by a black line and the radial critical curve and caustic are shown with a white line.

Table G1. The Bayesian evidence, ln $\mathcal{Z}$, of each model fit performed by the Mass pipelines using (i) a decomposed mass model assuming two Sersic profiles, an elliptical NFW and external shear; or (ii) a BPL mass model with external shear. Both models have the priors on various parameters adjusted such that they have a shallower inner density and can form a large radial critical curve. Log evidences are compared to the values found in the main paper, for models including an SMBH. Fits to both the F390W and F814W images are shown, where the F390W fits assume the Sersic parameters of the F814W image for the stellar mass. The favoured model is always that with an SMBH, because models with a shallow inner density fail to reconstruct the counter image's structure (see Fig. G2).

Filter	Model	Shallow Density	SMBH
F390W	Decomposed	125649.72	125 699.06
F390W	BPL	125548.22	125 693.78
F814W	Decomposed	78289.00	78 332.19
F814W	BPL	78238.79	78 329.28

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/521/3/3298/7085506 by guest on 07 April 2023